

QY	61	TGGAGTCCAGCAGGCCCTTCAGAGCTACCAGCAAGCCCTGCTCCGCACTCTCCCTAGACAAA	120
Db	61	TGGAGTCCAGCAGGCCCTTCAGAGCTACCAGCAAGCCCTGCTCCGCACTCTCCCTAGACAAA	120
QY	121	GTCCAGCGCAGCCTGGGCCCCCGAGCACCACCCAGCCTCCGAGGCATGTCCTCATCCATAAC	180
Db	121	GTCCAGCGCAGCCTGGGCCCCCGAGCACCACCCAGCCTCCGAGGCATGTCCTCATCCATAAC	180
QY	181	ACCCTCAAACAGCTGCAGGTGCACCTTCGCCTGGCTCCGCCCCCTGCCCTGCCCCCGAG	240
Db	181	ACCCTCAAACAGCTGCAGGTGCACCTTCGCCTGGCTCCGCCCCCTGCCCTGCCCCCGAG	240
QY	241	CCCCTCTTCCTGGCGAGGAGGATTCTCCCTGTTCAGCCACCATTTGGCTCTATCCTCAGG	300
Db	241	CCCCTCTTCCTGGCGAGGAGGATTCTCCCTGTTCAGCCACCATTTGGCTCTATCCTCAGG	300
QY	301	GAGCTGGACACCTCCATGGATGGGACTGAGCCCCCTCAGAAATCCAGTGACTCCCTTTGGC	360
Db	301	GAGCTGGACACCTCCATGGATGGGACTGAGCCCCCTCAGAAATCCAGTGACTCCCTTTGGC	360
QY	361	CTCCAGAAATGAAGTGCCACCCAGCCTGATCCAGTCTTCTTAGAAGCTCTGAGTCCCGG	420
Db	361	CTCCAGAAATGAAGTGCCACCCAGCCTGATCCAGTCTTCTTAGAAGCTCTGAGTCCCGG	420
QY	421	TACTTTGGGGACTCTGGCCTGGATGACTTCTTTCTGGACATTGACACATCTGCGGTAGAA	480
Db	421	TACTTTGGGGACTCTGGCCTGGATGACTTCTTTCTGGACATTGACACATCTGCGGTAGAA	480
QY	481	AAGGAGCCTGCACGGGCCCCACAGAGCCTCCTCACAACTTCTTGTGCCCCAGGTTCT	540
Db	481	AAGGAGCCTGCACGGGCCCCACAGAGCCTCCTCACAACTTCTTGTGCCCCAGGTTCT	540
QY	541	TGGGAGTGAATGAACCTGGATCACATCATGGAATCATTTCTGGGTCTTAA	591
Db	541	TGGGAGTGAATGAACCTGGATCACATCATGGAATCATTTCTGGGTCTTAA	591
RESULT 2			
CN304585			
LOCUS	CN304585	729 bp mRNA	linear EST 16-MAY-2004
DEFINITION	17000600512783 GRN_PRENEU Homo sapiens cDNA 5', mRNA sequence.		
ACCESSION	CN304585		
VERSION	CN304585.1	GI:47320999	
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE			
AUTHORS	Brandenberger,R., Wei,H., Zhang,S., Lei,S., Murage,J., Fisk,G.J., Li,Y., Xu,C., Fang,R., Guegler,K., Rao,M.S., Mandalam,R., Lebkowski,J and Stanton,L.W.		
TITLE	Transcriptome characterization elucidates signaling networks that control human ES cell growth and differentiation		
JOURNAL	Nat. Biotechnol. 22 (6), 707-716 (2004)		
COMMENT	Contact: Brandenberger R Regenerative Medicine Geron Corporation 230 Constitution Drive, Menlo Park, CA 94025, USA Tel: 650 473 8658 Fax: 650 473 7760 Email: rbrandenberger@geron.com Insert Length: 729 Std Error: 0.00. Location/Qualifiers		
FEATURES			
source	1. .729		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/tissue_type="embryonic stem cell, retinoic acid and mitogen-treated hES cell line H7"		
	/clone_lib="GRN_PRENEU"		
	/note="oligo dt primed, full-length enriched cDNA library		

ORIGIN			
Query Match	99.7%	Score 589.4; DB 7; Length 729;	
Best Local Similarity	99.8%;	Pred. No. 1.7e-135;	
Matches 590; Conservative	0; Mismatches	1; Indels	0; Gaps 0;
QY	1	ATGGAGGGAGGCTTGAAGAGGAAACACTCTGATTTTGGAAAGAGGAGGAGGAGGAGGTGGGAG	60
Db	98	ATGGTGGGAGGCTTGAAGAGGAAACACTCTGATTTTGGAAAGAGGAGGAGGAGGAGGTGGGAG	157
QY	61	TGGAGTCCAGCAGGCCTTCAGAGCTACCCAGCAAGCCCTGCTCCGCACTCTCCCTAGACAAA	120
Db	158	TGGAGTCCAGCAGGCCTTCAGAGCTACCCAGCAAGCCCTGCTCCGCACTCTCCCTAGACAAA	217
QY	121	GTCCAGCGCAGCCTGGGCCCCCGAGCACCCAGCCTCCGAGGCATGTCTCATCCATAAC	180
Db	218	GTCCAGCGCAGCCTGGGCCCCCGAGCACCCAGCCTCCGAGGCATGTCTCATCCATAAC	277
QY	181	ACCCTCAAACAGCTGCAGGCTGCACCTTCGCCTGGCTCCGCCCCCTGCCCTGCCCCCGAG	240
Db	278	ACCCTCAAACAGCTGCAGGCTGCACCTTCGCCTGGCTCCGCCCCCTGCCCTGCCCCCGAG	337
QY	241	CCCCTCTTCCTGGCGAGGAGGATTTCTCCCTGTTCAGCCACCATTTGGCTCTATCCTCAGG	300
Db	338	CCCCTCTTCCTGGCGAGGAGGATTTCTCCCTGTTCAGCCACCATTTGGCTCTATCCTCAGG	397
QY	301	GAGCTGGACACCTCCATGGATGGGACTGAGCCCCCTCAGAAATCCAGTGACTCCCTTTGGC	360
Db	398	GAGCTGGACACCTCCATGGATGGGACTGAGCCCCCTCAGAAATCCAGTGACTCCCTTTGGC	457
QY	361	CTCCAGAAATGAAGTGCCACCCAGCCTGATCCAGTCTTCTTAGAAGCTCTGAGTCCCGG	420
Db	458	CTCCAGAAATGAAGTGCCACCCAGCCTGATCCAGTCTTCTTAGAAGCTCTGAGTCCCGG	517
QY	421	TACTTTGGGGACTCTGGCCTGGATGACTTCTTTCTGGACATTGACACATCTGCGGTAGAA	480
Db	518	TACTTTGGGGACTCTGGCCTGGATGACTTCTTTCTGGACATTGACACATCTGCGGTAGAA	577
QY	481	AAGGAGCCTGCACGGGCCCCACAGAGCCTCCTCACAACTTCTTGTGCCCCAGGTTCT	540
Db	578	AAGGAGCCTGCACGGGCCCCACAGAGCCTCCTCACAACTTCTTGTGCCCCAGGTTCT	637
QY	541	TGGGAGTGAATGAACCTGGATCACATCATGGAATCATTTCTGGGTCTTAA	591
Db	638	TGGGAGTGAATGAACCTGGATCACATCATGGAATCATTTCTGGGTCTTAA	688
RESULT 3			
CN304584			
LOCUS	CN304584	742 bp mRNA	linear EST 16-MAY-2004
DEFINITION	17000597971761 GRN_PREHEP Homo sapiens cDNA 5', mRNA sequence.		
ACCESSION	CN304584		
VERSION	CN304584.1	GI:47320998	
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE			
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
TITLE	1 (bases 1 to 742)		
JOURNAL	Brandenberger,R., Wei,H., Zhang,S., Lei,S., Murage,J., Fisk,G.J., Li,Y., Xu,C., Fang,R., Guegler,K., Rao,M.S., Mandalam,R., Lebkowski,J and Stanton,L.W.		
COMMENT	Transcriptome characterization elucidates signaling networks that control human ES cell growth and differentiation Nat. Biotechnol. 22 (6), 707-716 (2004) Contact: Brandenberger R Regenerative Medicine Geron Corporation 230 Constitution Drive, Menlo Park, CA 94025, USA Tel: 650 473 8658		

Fax: 650 473 7760
Email: rbrandenberger@geron.com
Insert Length: 742 Std Error: 0.00.
Location/Qualifiers
source
1. .742
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="embryonic stem cells, DMSO-treated H9 cell line"
/clone_lib="GRN PREHEP"
/note="oligo dT primed, full-length enriched cDNA library from DMSO-treated hES cell line H9 (p22) maintained in feeder-free conditions"

ORIGIN
Query Match 99.7%; Score 589.4; DB 7; Length 742;
Best Local Similarity 99.8%; Pred. No. 1.7e-135;
Matches 590; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGAGGGAGGCTTGAAGAGGAAACACTCTGATTTGGAAGAGGAGGAGGAGGTGGAG 60
| | | | |
Db 68 ATGGTGGAGGCTTGAAGAGGAAACACTCTGATTTGGAAGAGGAGGAGGAGGTGGAG 127
| | | | |
QY 61 TGGAGTCCAGCAGGCTTTCAGAGCTACCAAGCAAGCCCTGCTCCGCATCTCCCTAGACAAA 120
| | | | |
Db 128 TGGAGTCCAGCAGGCTTTCAGAGCTACCAAGCAAGCCCTGCTCCGCATCTCCCTAGACAAA 187
| | | | |
QY 121 GTCCAGCGCAGCCTGGGCCCCGAGCACCCAGCCTCCGAGGCATGTCTCATCCATAAC 180
| | | | |
Db 188 GTCCAGCGCAGCCTGGGCCCCGAGCACCCAGCCTCCGAGGCATGTCTCATCCATAAC 247
| | | | |
QY 181 ACCCTCCAACAGCTGCAGGCTGCACATTGCGCTGGCTCCCGCCCTGCCCTGCCCCCGAG 240
| | | | |
Db 248 ACCCTCCAACAGCTGCAGGCTGCACATTGCGCTGGCTCCCGCCCTGCCCTGCCCCCGAG 307
| | | | |
QY 241 CCCCTCTTCTGGCGGAGGAGGATTTCTCCCTGTCCAGCCACCATTTGGCTCTATCCTCAGG 300
| | | | |
Db 308 CCCCTCTTCTGGCGGAGGAGGATTTCTCCCTGTCCAGCCACCATTTGGCTCTATCCTCAGG 367
| | | | |
QY 301 GAGCTGGACACCTCCATGGATGGGACTGAGCCCCCTCAGAAATCCAGTACTCCCCCTTGGC 360
| | | | |
Db 368 GAGCTGGACACCTCCATGGATGGGACTGAGCCCCCTCAGAAATCCAGTACTCCCCCTTGGC 427
| | | | |
QY 361 CTCCAGAAATGAAGTGCCACCCAGCCTGATCCAGTCTTCTTAGAAGCTCTGAGCTCCCGG 420
| | | | |
Db 428 CTCCAGAAATGAAGTGCCACCCAGCCTGATCCAGTCTTCTTAGAAGCTCTGAGCTCCCGG 487
| | | | |
QY 421 TACTTGGGGACTCTGGCCTTGGATGACTTCTTTCTGGACATTGACACATCTGCGGTAGAA 480
| | | | |
Db 488 TACTTGGGGACTCTGGCCTTGGATGACTTCTTTCTGGACATTGACACATCTGCGGTAGAA 547
| | | | |
QY 481 AAGGAGCCTGCACGGGCCCCACAGAGCCTCCTCACAACCTTCTTGTGCCCCCAGGTTCT 540
| | | | |
Db 548 AAGGAGCCTGCACGGGCCCCACAGAGCCTCCTCACAACCTTCTTGTGCCCCCAGGTTCT 607
| | | | |
QY 541 TGGGAGTGAATGAACCTGGATCACATCATGGAAATCATTTGGGGTCTCTAA 591
| | | | |
Db 608 TGGGAGTGAATGAACCTGGATCACATCATGGAAATCATTTGGGGTCTCTAA 658
| | | | |

RESULT 4
BM560308
LOCUS
DEFINITION BM560308 1038 bp mRNA linear EST 20-FEB-2002
5', mRNA sequence.
AGENCOURT_6563953 NIH_MGC_119 Homo sapiens cDNA clone IMAGE:5744719
BM560308
ACCESSION BM560308.1 GI:18804629
VERSION
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
1 (bases 1 to 1038)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12766 row: i column: 08
High quality sequence start: 32
High quality sequence stop: 703.
Location/Qualifiers
source
1. .1038
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5744719"
/tissue_type="medulla"
/lab_host="DH10B"
/clone_lib="NIH_MGC_119"
/note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA source normal medulla from anonymous male age 27. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.3 kb, insert size range 0.9-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 013. Note: this is a NIH_MGC Library."

ORIGIN
Query Match 99.7%; Score 589.4; DB 4; Length 1038;
Best Local Similarity 99.8%; Pred. No. 1.8e-135;
Matches 590; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGAGGGAGGCTTGAAGAGGAAACACTCTGATTTGGAAGAGGAGGAGGAGGTGGAG 60
| | | | |
Db 100 ATGGTGGAGGCTTGAAGAGGAAACACTCTGATTTGGAAGAGGAGGAGGAGGTGGAG 159
| | | | |
QY 61 TGGAGTCCAGCAGGCTTTCAGAGCTACCAGCAAGCCCTGCTCCGCATCTCCCTAGACAAA 120
| | | | |
Db 160 TGGAGTCCAGCAGGCTTTCAGAGCTACCAGCAAGCCCTGCTCCGCATCTCCCTAGACAAA 219
| | | | |
QY 121 GTCCAGCGCAGCCTGGGCCCCCGAGCACCCAGCCTCCGAGGCATGTCTCATCCATAAC 180
| | | | |
Db 220 GTCCAGCGCAGCCTGGGCCCCCGAGCACCCAGCCTCCGAGGCATGTCTCATCCATAAC 279
| | | | |
QY 181 ACCCTCCAACAGCTGCAGGCTGCACCTTGCCTGGCTCCCGCCCTGCCCTGCCCCCGAG 240
| | | | |
Db 280 ACCCTCCAACAGCTGCAGGCTGCACCTTGCCTGGCTCCCGCCCTGCCCTGCCCCCGAG 339
| | | | |
QY 241 CCCCTCTTCTGGGCGAGGAGGATTTCTCCCTGTCCAGCCACCATTTGGCTCTATCCTCAGG 300
| | | | |
Db 340 CCCCTCTTCTGGGCGAGGAGGATTTCTCCCTGTCCAGCCACCATTTGGCTCTATCCTCAGG 399
| | | | |
QY 301 GAGCTGGACACCTCCAATGGATGGGACTGAGCCCCCTCAGAAATCCAGTACTCCCCCTTGGC 360
| | | | |
Db 400 GAGCTGGACACCTCCAATGGATGGGACTGAGCCCCCTCAGAAATCCAGTACTCCCCCTTGGC 459
| | | | |
QY 361 CTCCAGAAATGAAGTGCCACCCAGCCTGATCCAGTCTTCTTAGAAGCTCTGAGCTCCCGG 420
| | | | |
Db 460 CTCCAGAAATGAAGTGCCACCCAGCCTGATCCAGTCTTCTTAGAAGCTCTGAGCTCCCGG 519
| | | | |
QY 421 TACTTGGGGACTCTGGCCTTGGATGACTTCTTTCTGGACATTTGACACATCTGCGGTAGAA 480
| | | | |
Db 520 TACTTGGGGACTCTGGCCTTGGATGACTTCTTTCTGGACATTTGACACATCTGCGGTAGAA 579
| | | | |
QY 481 AAGGAGCCTGCACGGGCCCCACAGAGCCTCCTCACAACCTTCTTGTGCCCCCAGGTTCT 540
| | | | |

||||| 580 AAGGAGCCTGCACGGGCCCCACCAGAGCCTCCTCAACACCTCTTCTGTGCCCCAGGTTCT 639

QY 541 TGGGAGTGAATGAACCTGGATCACATCATGGAATCATTTCTGGGTCCTAA 591

Db 640 TGGGAGTGAATGAACCTGGATCACATCATGGAATCATTTCTGGGTCCTAA 690

RESULT 5

CR605181 1298 bp mRNA linear HTC 21-JUL-2004

LOCUS full-length cDNA clone CS0DJ012YJ05 of T cells (Jurkat cell line)

DEFINITION Cot 10-normalized of Homo sapiens (human).

ACCESSION CR605181

VERSION CR605181.1 GI:50485988

KEYWORDS HTC; CNSLT_cDNA.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 1298)

AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.

TITLE Full-length cDNA libraries and normalization

JOURNAL Unpublished

REMARK Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ InvitroGen Corporation 1600 Faraday Avenue

2 (bases 1 to 1298)

Genoscope.

AUTHORS Direct Submission

TITLE Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :

JOURNAL BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)

COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen.

FEATURES Location/Qualifiers

source 1..1298

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="CS0DJ012YJ05"

/tissue_type="T cells (Jurkat cell line) Cot 10-normalized"

/plasmid="pCMVSPORT_6"

ORIGIN

Query Match 99.7%; Score 589.4; DB 3; Length 1298;

Best Local Similarity 99.8%; Pred. No. 1.9e-135;

Matches 590; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGAGGGAGGCTTGAAGAGGAAACACTCTGATTTGGAAGAGGAGGAGGAGGTGGGAG 60

Db 105 ATGGTGGGAGGCTTGAAGAGGAAACACTCTGATTTGGAAGAGGAGGAGGAGGTGGGAG 164

QY 61 TGGAGTCCAGCAGGCCCTTCAGAGCTACCAGCAAGCCCTGCTCCGCATCTCCCTAGACAAA 120

Db 165 TGGAGTCCAGCAGGCCCTTCAGAGCTACCAGCAAGCCCTGCTCCGCATCTCCCTAGACAAA 224

QY 121 GTCCAGCGCAGCCTGGGCCCCCGAGCACCCAGCCTCCGCGAGGCATGTCCTATCCATAAC 180

Db 225 GTCCAGCGCAGCCTGGGCCCCCGAGCACCCAGCCTCCGCGAGGCATGTCCTATCCATAAC 284

QY 181 ACCCTCCAACAGCTGCAGGCTGCACTTTCGCCTGGCTCCCGCCCTGCCCTGCCCCCGGAG 240

Db 285 ACCCTCCAACAGCTGCAGGCTGCACTTTCGCCTGGCTCCCGCCCTGCCCTGCCCCCGGAG 344

QY 241 CCCCTCTTCTGGCGGAGGAGGATTTCTCCCTGTTCAGCCACCATTGGCTCTATCCTCAGG 300

Db 345 CCCCTCTTCTGGCGGAGGAGGATTTCTCCCTGTTCAGCCACCATTGGCTCTATCCTCAGG 404

QY 301 GAGCTGGACACCTCCATGGATGGGACTGAGCCCCCTCAGAATCCAGTGAATCCCTTGGC 360

Db 405 GAGCTGGACACCTCCATGGATGGGACTGAGCCCCCTCAGAATCCAGTGAATCCCTTGGC 464

QY 361 CTCCAGAATGAAGTGCACCCAGCCCTGATCCAGTCTTCTTAGAAGCTCTGAGCTCCCGG 420

Db 465 CTCCAGAATGAAGTGCACCCAGCCCTGATCCAGTCTTCTTAGAAGCTCTGAGCTCCCGG 524

QY 421 TACTTGGGGGACTCTGGCCTGGATGACTTCTTCTTGGACATTGACACATCTGCGGTAGAA 480

Db 525 TACTTGGGGGACTCTGGCCTGGATGACTTCTTCTTGGACATTGACACATCTGCGGTAGAA 584

QY 481 AAGGAGCCTGCACGGGCCCCCACCAGAGCCTCCTTCACAACCTCTTCTGTGCCAGGTTCT 540

Db 585 AAGGAGCCTGCACGGGCCCCCACCAGAGCCTCCTTCACAACCTCTTCTGTGCCAGGTTCT 644

QY 541 TGGGAGTGAATGAAGTGGATCACATCATATGGAATCATTTCTGGGGTCCTAA 591

Db 645 TGGGAGTGAATGAAGTGGATCACATCATATGGAATCATTTCTGGGGTCCTAA 695

RESULT 6

CR601245 1318 bp mRNA linear HTC 21-JUL-2004

LOCUS full-length cDNA clone CS0DI057YM14 of Placenta Cot 25-normalized

DEFINITION of Homo sapiens (human).

ACCESSION CR601245

VERSION CR601245.1 GI:50482052

KEYWORDS HTC; CNSLT_cDNA.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 1318)

AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.

TITLE Full-length cDNA libraries and normalization

JOURNAL Unpublished

REMARK Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ InvitroGen Corporation 1600 Faraday Avenue

2 (bases 1 to 1318)

Genoscope.

AUTHORS Direct Submission

TITLE Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :

JOURNAL BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)

COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen.

FEATURES Location/Qualifiers

source 1..1318

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="CS0DI057YM14"

/tissue_type="Placenta Cot 25-normalized"

/plasmid="pCMVSPORT_6"

ORIGIN

Query Match 99.7%; Score 589.4; DB 3; Length 1318;

Best Local Similarity 99.8%; Pred. No. 1.9e-135;

Matches 590; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGAGGGAGGCTTGAAGAGGAAACACTCTGATTTGGAAGAGGAGGAGGAGGTGGGAG 60

Db 112 ATGGTGGGAGGCTTGAAGAGGAAACACTCTGATTTGGAAGAGGAGGAGGAGGTGGGAG 171

QY 61 TGGAGTCCAGCAGGCCCTTCAGAGCTACCAGCAAGCCCTGCTCCGCATCTCCCTAGACAAA 120

Db 172 TGGAGTCCAGCAGGCCCTTCAGAGCTACCAGCAAGCCCTGCTCCGCATCTCCCTAGACAAA 231

QY 121 GTCCAGCGCAGCCTGGGCCCCCGAGCACCAGCCTCCGAGGCGATGCTCCTCATCCATAAC 180
Db 232 GTCCAGCGCAGCCTGGGCCCCCGAGCACCAGCCTCCGAGGCGATGCTCCTCATCCATAAC 291
QY 181 ACCCTCCAACAGCTGCAGGTGCACCTTGGCTGGCTCCGCCCCCTGCCCTGCCCTCCCGAG 240
Db 292 ACCCTCCAACAGCTGCAGGTGCACCTTGGCTGGCTCCGCCCCCTGCCCTGCCCTCCCGAG 351
QY 241 CCCCTCTTCTGGGCGAGGAGGATTTCTCCCTGTGAGCCACCATTTGGCTCTATCTCAGG 300
Db 352 CCCCTCTTCTGGGCGAGGAGGATTTCTCCCTGTGAGCCACCATTTGGCTCTATCTCAGG 411
QY 301 GAGCTGGACACCTCCATGGATGGGACTGAGCCCCCTCAGAATCCAGTACTCCCTTGGC 360
Db 412 GAGCTGGACACCTCCATGGATGGGACTGAGCCCCCTCAGAATCCAGTACTCCCTTGGC 471
QY 361 CTCCAGAAATGAAGTGCCACCCAGCCTGATCCAGTCTTTAGAACTCTGAGCTCCCGG 420
Db 472 CTCCAGAAATGAAGTGCCACCCAGCCTGATCCAGTCTTTAGAACTCTGAGCTCCCGG 531
QY 421 TACTTGGGGACTCTGGCCTGGATGACTTCTTTCTGGACATTTGACACATCTGCGGTAGAA 480
Db 532 TACTTGGGGACTCTGGCCTGGATGACTTCTTTCTGGACATTTGACACATCTGCGGTAGAA 591
QY 481 AAGGAGCCTGCACGGGCCCCCAGAGCCTCCTCAACACCTCTTCTGTGCCCCCAGGTTCT 540
Db 592 AAGGAGCCTGCACGGGCCCCCAGAGCCTCCTCAACACCTCTTCTGTGCCCCCAGGTTCT 651
QY 541 TGGGAGTGAATGAATGATGATCAGATCATGGAATCAATTTCTGGGTCTCTAA 591
Db 652 TGGGAGTGAATGAATGATGATCAGATCATGGAATCAATTTCTGGGTCTCTAA 702

RESULT 7
BU157342 1056 bp mRNA linear EST 04-SEP-2002
LOCUS AGENCOURT_6925371 NIH_MGC_110 Homo sapiens cDNA clone IMAGE:5952523
DEFINITION 5', mRNA sequence.
ACCESSION BU157342 GI:22670874
VERSION BU157342.1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE NIH-MGC <http://mgi.nci.nih.gov/>.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgaps-r@mail.nih.gov
Tissue Procurement: ATCC
cdna Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LCM2140 row: k column: 20
High quality sequence stop: 587.
Location/Qualifiers

FEATURES
source 1..1056
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5952523"
/tissue_type="ductal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_110"
/note="Organ: pancreas; Vector: pOTB7; Site 1: XhoI;
Site 2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed by

ORIGIN

Query Match 98.6%; Score 582.6; DB 5; Length 1056;
Best Local Similarity 99.0%; Pred. No. 8.8e-134;
Matches 585; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 1 ATGGAGGGAGGCTTGAAGAGGAAACACTCTGATTTGGAAAGAGGAGGAGGAGGTGGAG 60
Db 108 ATGGTGGGAGGCTTGAAGAGGAAACACTCTGATTTGGAAAGAGGAGGAGGAGGTGGAG 167
QY 61 TGGAGTCCAGCAGGCTTCCAGAGCTACCAAGCAAGCCCTGCTCCGATCTCCCTAGACAAA 120
Db 168 TGGAGTCCAGCAGGCTTCCAGAGCTACCAAGCAAGCCCTGCTCCGATCTCCCTAGACAAA 227
QY 121 GTCCAGCGCAGCCTGGGCCCCCGAGCACCCAGCCTCCGAGGCGATGCTCCTCATCCATAAC 180
Db 228 GTCCAGCGCAGCCTGGGCCCCCGAGCACCCAGCCTCCGAGGCGATGCTCCTCATCCATAAC 287
QY 181 ACCCTCCAACAGCTGCAGGCTGCACCTTCGCTGGCTCCGCCCCCTGCCCTGCCCTCCCGAG 240
Db 288 ACCCTCCAACAGCTGCAGGCTGCACCTTCGCTGGCTCCGCCCCCTGCCCTGCCCTCCCGAG 347
QY 241 CCCCTCTTCTGGGCGAGGAGGATTTCTCCCTGTGAGCCACCATTTGGCTCTATCTCAGG 300
Db 348 CCCCTCTTCTGGGCGAGGAGGATTTCTCCCTGTGAGCCACCATTTGGCTCTATCTCAGG 407
QY 301 GAGCTGGACACCTCCATGGATGGGACTGAGCCCCCTCAGAATCCAGTACTCCCTTGGC 360
Db 408 GAGCTGGACACCTCCATGGATGGGACTGAGCCCCCTCAGAATCCAGTACTCCCTTGGC 467
QY 361 CTCCAGAAATGAAGTGCCACCCAGCCTGATCCAGTCTTTAGAACTCTGAGCTCCCGG 420
Db 468 CTCCAGAAATGAAGTGCCACCCAGCCTGATCCAGTCTTTAGAACTCTGAGCTCCCGG 527
QY 421 TACTTGGGGACTCTGGCCTGGATGACTTCTTTCTGGACATTTGACACATCTGCGGTAGAA 480
Db 528 TACTTGGGGACTCTGGCCTGNATGACTTCTTTCTGGACATTTGACACATCTGCGGTAGAA 587
QY 481 AAGGAGCCTGCACGGGCCCCCAGAGCCTCCTCAACACCTCTTCTGTGCCCCCAGGTTCT 540
Db 588 AAGGAGCCTGCACGGGCCCCCAGAGCCTCCTCAACACCTCTTCTGTGCCCCCAGGTTCT 647
QY 541 TGGGAGTGAATGAATGATGATCAGATCATGGAATCAATTTCTGGGTCTCTAA 591
Db 648 TGGGAGTGAATGAATGATGATCAGATCATGGAATCAATTTCTGGGTCTCTAA 698

RESULT 8
CN304590
LOCUS 17000600179006 GRN_PRENEU Homo sapiens cDNA 5', mRNA sequence.
DEFINITION CN304590
ACCESSION CN304590.1 GI:47321004
VERSION EST.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE 1 (bases 1 to 646)
JOURNAL Brandenberger, R., Wei, H., Zhang, S., Lei, S., Murage, J., Fisk, G.J.,
COMMENT Li, Y., Xu, C., Fang, R., Guegler, K., Rao, M.S., Mandalam, R.,
Lebkowski, J. and Stanton, L.W.
Transcriptome characterization elucidates signaling networks that
control human ES cell growth and differentiation
Nat. Biotechnol. 22 (6), 707-716 (2004)
Contact: Brandenberger R
Regenerative Medicine
Geron Corporation
230 Constitution Drive, Menlo Park, CA 94025, USA

Tel: 650 473 8658
Fax: 650 473 7760
Email: rbrandenberger@geron.com
Insert Length: 646 Std Error: 0.00.
Location/Qualifiers
1. .646
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue_type="embryonic stem cell, retinoic acid and
mitogen-treated hES cell line H7"
/clone_lib="GRN PRENEU"
/note="Oligo dT⁻primed, full-length enriched cDNA library
from hES cell line H7 (p29) maintained in feeder-free
conditions. Embryoid bodies were generated in the presence
of all-trans retinoic and mitogens."

ORIGIN

Query Match 98.2%; Score 580.4; DB 7; Length 646;
Best Local Similarity 99.8%; Pred. No. 2.8e-133;
Matches 581; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGAGGGAGGCTTGAAGAGGAAACACTCTGATTTGGAAGAGGAGGAGGAGGTGGGAG 60
Db 65 ATGGTGGAGGCTTGAAGAGGAAACACTCTGATTTGGAAGAGGAGGAGGAGGTGGGAG 124

QY 61 TGGAGTCCAGCAGGCCCTTCAGAGCTACCAGCAAGCCCTGCTCCGCATCTCCCTAGACAAA 120
Db 125 TGGAGTCCAGCAGGCCCTTCAGAGCTACCAGCAAGCCCTGCTCCGCATCTCCCTAGACAAA 184

QY 121 GTCCAGCGCAGCCTGGGCCCCCGAGCACCAGCCCTCCGCAGGCATGTCCTCATCCATAAC 180
Db 185 GTCCAGCGCAGCCTGGGCCCCCGAGCACCAGCCCTCCGCAGGCATGTCCTCATCCATAAC 244

QY 181 ACCCTCCAACAGCTGCAGGCTGCACCTTCGCCTGGCTCCCGCCCCCTGCCCTGCCCTCCCGAG 240
Db 245 ACCCTCCAACAGCTGCAGGCTGCACCTTCGCCTGGCTCCCGCCCCCTGCCCTGCCCTCCCGAG 304

QY 241 CCCCTCTTCTGGCGAGGAGGATTCTCCCTGTGACGCCACCATTTGGCTCTATCCTCAGG 300
Db 305 CCCCTCTTCTGGCGAGGAGGATTCTCCCTGTGAGCCACCATTGGCTCTATCCTCAGG 364

QY 301 GAGCTGGACACCTCCATGGATGGGACTGAGCCCCCTCAGAATCCAGTGACTCCCTTGGC 360
Db 365 GAGCTGGACACCTCCATGGATGGGACTGAGCCCCCTCAGAATCCAGTGACTCCCTTGGC 424

QY 361 CTCCAGAAATGAAGTGCACCCCGCCAGCCTGATCCAGTCTTCTTAGAAGCTCTGAGTCCCGG 420
Db 425 CTCCAGAAATGAAGTGCACCCCGCCAGCCTGATCCAGTCTTCTTAGAAGCTCTGAGTCCCGG 484

QY 421 TACTTGGGGACTCTGGCCTGGATGACTTCTTTCTGGACATTGACACATCTGCGGTAGAA 480
Db 485 TACTTGGGGACTCTGGCCTGGATGACTTCTTTCTGGACATTGACACATCTGCGGTAGAA 544

QY 481 AAGGAGCCTGCACGGGCCCCACAGAGCCTCCTCAACAACCTCTTCTGTGCCCGAGGTTCT 540
Db 545 AAGGAGCCTGCACGGGCCCCACAGAGCCTCCTCAACAACCTCTTCTGTGCCCGAGGTTCT 604

QY 541 TGGGAGTGAATGAACCTGGATCACATCATGGAAATCATTTCTG 582
Db 605 TGGGAGTGAATGAACCTGGATCACATCATGGAAATCATTTCTG 646

RESULT 9
BX398356 1017 bp mRNA linear EST 28-APR-2004
LOCUS BX398356 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
DEFINITION clone CS0DI057YM14 5-PRIME, mRNA sequence.
ACCESSION BX398356
VERSION BX398356.2 GI:46846521
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1017)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On May 13, 2003 this sequence version replaced gi:30613656.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
9131.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CS0DI057BG07QPI&c=9131.r.

FEATURES
source
1. .1017
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DI057YM14"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and EcoR V sites of the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 97.4%; Score 575.8; DB 5; Length 1017;
Best Local Similarity 99.5%; Pred. No. 4.3e-132;
Matches 588; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 1 ATGGAGGGAGGCTTGAAGAGGAAACACTCTGATTTGGAAGAGGAGGAGGAGGTGGGAG 60
Db 112 ATGGTGGAGGCTTGAAGAGGAAACACTCTGATTTGGAAGAGGAGGAGGAGGTGGGAG 171

QY 61 TGGAGTCCAGCAGGCCCTTCAGAGCTACCAGCAAGCCCTGCTCCGCATCTCCCTAGACAAA 120
Db 172 TGGAGTCCAGCAGGCCCTTCAGAGCTACCAGCAAGCCCTGCTCCGCATCTCCCTAGACAAA 231

QY 121 GTCCAGCGCAGCTGGGCCCCCGAGCACCAGCCTCCGAGGCATGTCCTCATCCATAAC 180
Db 232 GTCCAGCGCA-CCTGGGCCCCCGAGCACCAGCCTCCGAGGCATGTCCTCATCCATAAC 290

QY 181 ACCCTCCAACAGCTGCAGGCTGCACCTTCGCCTGGCTCCCGCCCCCTGCCCTCCCGAG 240
Db 291 ACCCTCCAACAGCTGCAGGCTGCACCTTCGCCTGGCTCCCGCCCCCTGCCCTCCCGAG 350

QY 241 CCCCTCTTCTGGCGAGGAGGATTCTCCCTGTGAGCCACCATTGGCTCTATCCTCAGG 300
Db 351 CCCCTCTTCTGGCGAGGAGGATTCTCCCTGTGAGCCACCATTGGCTCTATCCTCAGG 410

QY 301 GAGCTGGACACCTCCATGGATGGGACTGAGCCCCCTCAGAAATCCAGTGACTCCCTTGGC 360
Db 411 GAGCTGGACACCTCCATGGATGGGACTGAGCCCCCTCAGAAATCCAGTGACTCCCTTGGC 470

QY 361 CTCCAGAAATGAAGTGCCACCCAGCCTGATCCAGTCTTCTTAGAAGCTCTGAGCTCCCGG 420
Db 471 CTCCAGAAATGAAGTGCCACCCAGCCTGATCCAGTCTTCTTAGAAGCTCTGAGCTCCCGG 530

QY 421 TACTTGGGGACTCTGGCCTGGATGACTTCTTTCTGGACATTGACACATCTGCGGTAGAA 480
Db 531 TACTTGGGGACTCTGGCCTGGATGACTTCTTTCTGGACATTGACACATCTGCGGTAGAA 590

QY 481 AAGGAGCCTGCACGGGCCCCACAGAGCCTCCTCAACAACCTCTTCTGTGCCCGAGGTTCT 540
Db 591 AAGGAGCCTGCACGGGCCCCACAGAGCCTCCTCAACAACCTCTTCTGTGCCCGAGGTTCT 650

QY 361 CTCCAGAAATGAAGTGCACCCAGCCTGATCCAGTCTTCTTAGAAGCTCTGAGCTCCCGG 420
|||||
Db 464 CTCCAGAAATGAAGTGCACCCAGCCTGATCCAGTCTTCTTAGAAGCTCTGAGCTCCCGG 523
|||||
QY 421 TACTTGGGGACTCTGGCTGGATGACTTCTTTCTGGACATTGACACATCTGCGGTAGAA 480
|||||
Db 524 TACTTGGGGACTCTGGCTGGATGACTTCTTTCTGGACATTGACACATCTGCGGTAGAA 583
|||||
QY 481 AAGGAGCCTGCACGGGCCACACAGAGCCTCCTCAACAACCTCTTCTGTGCCCCAGGTTCT 540
|||||
Db 584 AAGGAGCCTGCACGGGCCACACAGAGCCTCCTCAACAACCTCTTCTGTGCCCCAGGTTCT 643
|||||
QY 541 TG-GGAGTGAATGAACCTGGATCACATCATGGAATCATTTCTGGGG 585
|||||
Db 644 TGGGAGTGAATGAACCTGGATCACATCATGGAATCATTTCTCGGG 689
|||||

RESULT 15
BM546306
LOCUS BM546306 1045 bp mRNA linear EST 20-FEB-2002
DEFINITION AGENCOURT 6498625 NIH_MGC_125 Homo sapiens cDNA clone IMAGE:5722856
5', mRNA Sequence.
ACCESSION BM546306
VERSION BM546306.1 GI:18779156
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1045)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Invitrogen
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM12709 row: j column: 09
High quality sequence start: 20
High quality sequence stop: 626.
Location/Qualifiers

FEATURES
source 1..1045
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5722856"
/lab_host="DH10B"
/clone_lib="NIH_MGC_125"
/note="Organ: ovary (pool of 3); Vector: pCMV-SPORT6;
Site_1: EcoRV (destroyed); Site_2: NotI; RNA source pool
of three ovaries, from females ranging in age from 38 to
49 yo. Library is oligo-dr primed and directionally cloned
(EcoRV site is destroyed upon cloning). Average insert
size 2.1 kb, insert size range 1-3.5 kb. Library is
normalized and enriched for full-length clones and was
constructed by C. Gruber (Invitrogen). Research Genetics
tracking code 036."

ORIGIN

Query Match 94.2%; Score 557; DB 4; Length 1045;
Best Local Similarity 98.6%; Pred. No. 2e-127;
Matches 582; Conservative 0; Mismatches 6; Indels 2; Gaps 2;
QY 1 ATGGAGGGAGGCTTGAAGAGGAAACACTCTGATTGGAAGAGGAGGAGGAGGTTGGGAG 60
|||||
Db 208 ATGGTGGGAGGCTTGAAGAGGAAACACTCTGATTGGAAGAGGAGGAGGAGGTTGGGAG 267
|||||
QY 61 TGGAGTCCAGCAGGCCTTCAGAGCTACCAGCAAGCCCTGCTCCGCATCTCCCTAGACAAA 120
|||||

Db 268 TGGAGTCCAGCAGGCCTTCAGAGCTACCAGCAAGCCCTGCTCCGCATCTCCCTAGACAAA 327
|||||
QY 121 GTCCAGCGCAGCCTTGGGCCCCCGAGCACCCAGCCTCCGAGGCATGTCTCTCATCCATAAC 180
|||||
Db 328 GTCCAGCGCAGCCTTGGGCCCCCGAGCACCCAGCCTCCGAGGCATGTCTCTCATCCATAAC 387
|||||
QY 181 ACCCTCCAACAGCTGCAGGCTGCATTGCGCTGGCTCCGCCCCCTGCCCTGCCCTCCCGGAG 240
|||||
Db 388 ACCCTCCAACAGCTGCAGGCTGCATTGCGCTGGCTCCGCCCCCTGCCCTGCCCTCCCGGAG 447
|||||
QY 241 CCCCTCTTCTTGGCGAGGAGGATTTCTCCCTGTGAGCCACCATTTGGCTCTATCTCAGG 300
|||||
Db 448 CCCCTCTTCTTGGCGAGGAGGATTTCTCCCTGTGAGCCACCATTTGGCTCTATCTCAGG 507
|||||
QY 301 GAGCTGGACACCTCCATGGATGGGACTGAGCCCCCTCAGAAATCCAGTCACTCCCTTGGC 360
|||||
Db 508 GAGCTGGACACCTCCATGGATGGGACTGAGCCCCCTCANAATCCAGTCACTCCCTTGGC 567
|||||
QY 361 CTCCAGAAATGAAGTGCACCCAGCCTGATCCAGTCTTCTTAGAAGCTCTGAGCTCCCGG 420
|||||
Db 568 CTCCAGAAATGAAGTGCACCCAGCCTGATCCAGTCTTCTTAGAAGCTCTGAGCTCCCGG 627
|||||
QY 421 TACTTGGGGGACTCTGGCCTGGATGACTTTCTTTCTGGACATTGACACATCTGCGGTAGAA 480
|||||
Db 628 TACTTGGGGGACTCTGGCCTGGATGACTTTCTTTCTGGACATTGACACATCTGCGGTAGAA 687
|||||
QY 481 AAGGAGCCTGCACGGGCCCCACACAGAGCCTCCTCAACAACCTCTTCTGTGCCCTCAGGTTCT 540
|||||
Db 688 AAGGAGCCTGCACGGGCCCCACCAAGCCTCCTCAACAACCTCTTCTGTGCCCTCAGGTTCT 747
|||||
QY 541 TGGGAGTGG-AATGAACTGGATCACATCAT-GGAAATCATTTCTGGGGTCC 588
|||||
Db 748 TGGGAGTGGCAATGACCTGGATCCCCCTCATGGGAAATCATTTCTGGGGTCC 797
|||||

Search completed: April 26, 2005, 10:11:12
Job time : 3125 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model
Run on: April 25, 2005, 12:21:55 ; Search time 166 Seconds
(without alignments)
456.657 Million cell updates/sec

Title: US-10-069-386A-2
Perfect score: 1033
Sequence: 1 MEGGLKRRKHSLEEEERWE.....APGSWEWNELDHIMEILGS 196

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
1: Geneseqp1980a:*
2: Geneseqp1990a:*
3: Geneseqp2000a:*
4: Geneseqp2001a:*
5: Geneseqp2002a:*
6: Geneseqp2003a:*
7: Geneseqp2003ba:*
8: Geneseqp2004a:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1033	100.0	196	4	AAB35402	Aab35402 Replicati
2	1026	99.3	196	4	AAM93922	Aam93922 Human pol
3	1026	99.3	196	8	ADL32052	Adl32052 Human pro
4	791	76.6	157	4	AAG74742	Aag74742 Human col
5	645.5	62.5	142	8	ABO60075	AbO60075 Human gen
6	402	38.9	111	8	ABO60074	AbO60074 Human gen
7	216	20.9	236	2	AAy36004	Aay36004 Extended
8	216	20.9	236	8	ADP19312	Adp19312 Human sec
9	213	20.6	236	3	AAy44362	Aay44362 Human cel
10	213	20.6	236	4	AAM93724	Aam93724 Human pol
11	213	20.6	236	4	AAG89292	Aag89292 Human sec
12	213	20.6	236	8	ADL31644	Adl31644 Human pro
13	213	20.6	278	4	AAM25550	Aam25550 Human pro
14	211	20.4	236	2	AAy31829	Aay31829 Human adu
15	205	19.8	222	8	ABO60367	AbO60367 Human gen
16	205	19.8	236	2	AAy02619	Aay02619 Amino aci
17	136	13.2	237	5	AAO22897	Aao22897 Mouse hae
18	125	12.1	314	5	ABP65076	Abp65076 Hypoxia-i
19	125	12.1	314	8	ADR14626	Adr14626 Human NF-
20	106	10.3	241	4	AAM39725	Aam39725 Human pol
21	106	10.3	254	3	AAB58258	Aab58258 Lung canc
22	106	10.3	254	4	AAG73682	Aag73682 Human col
23	106	10.3	254	4	AAM41511	Aam41511 Human pol
24	105.5	10.2	578	4	AAM38707	Aam38707 Human pol
25	105.5	10.2	620	7	ADN95361	Adn95361 Human BEC

26	105.5	10.2	620	8	ADP54705	Adp54705 Human PRO
27	105.5	10.2	650	4	AAM38706	Aam38706 Human pol
28	105.5	10.2	685	7	ADE28201	Ade28201 Human MDD
29	105.5	10.2	759	4	AAM40492	Aam40492 Human pol
30	105.5	10.2	759	4	AAM40493	Aam40493 Human pol
31	105.5	10.2	760	7	ADN95128	Adn95128 Human LEC
32	105	10.2	241	5	ABB06375	Abb06375 Human CHD
33	105	10.2	241	5	AAO22898	Aao22898 Human hae
34	105	10.2	241	8	ABM81593	Abm81593 Tumour-as
35	104.5	10.1	740	4	ABB11713	Abb11713 Human KIA
36	104.5	10.1	879	5	ABP69285	Abp69285 Human pol
37	104	10.1	322	4	ABG22868	Abg22868 Novel hum
38	101.5	9.8	728	7	ADD46841	Add46841 Human Pro
39	101.5	9.8	728	7	ADE56294	Ade56294 Human Pro
40	101.5	9.8	728	8	ADJ66562	Adj66562 PI3 kinas
41	98	9.5	562	7	ADB65060	Adb65060 Human pro
42	98	9.5	562	8	ADR58973	Adr58973 Human Elk
43	98	9.5	562	8	ADR58971	Adr58971 Human Elk
44	97	9.4	803	8	ADN19883	Adn19883 Bacterial
45	96	9.3	817	6	ABR53281	Abr53281 Protein s

ALIGNMENTS

RESULT 1
AAB35402
ID AAB35402 standard; protein; 196 AA.
XX
AC AAB35402;
XX
DT 23-MAY-2001 (first entry)
XX
DE Replication protein A binding transcriptional activator 1 RBT1.
XX
KW RBT1; replication protein A binding transcriptional activator 1 RBT1.
KW gene therapy; apoptosis; cancer; leukaemia.
XX
OS Unidentified.
XX
PN WO200114546-A2.
XX
PD 01-MAR-2001.
XX
PF 17-AUG-2000; 2000WO-CA000948.
XX
PR 19-AUG-1999; 99US-0149472P.
XX
(TRAN-) CENT TRANSLATIONAL RES IN CANCER.
PA Alaoui-Jamali MA, Cho JM;
XX
WPI; 2001-218447/22.
DR N-PSDB; AAF28052.
XX
PT Novel replication protein A binding transcriptional activator 1 gene,
PT useful for treating neoplastic disorders such as cancer and in gene
PT therapy.
XX
PS Disclosure; Fig 1; 16pp; English.
XX
CC The present invention provides the protein and coding sequences of the
CC replication protein A binding transcriptional activator 1 (RBT1). The
CC protein is capable of inducing apoptosis. The sequences are useful in the
CC gene therapy and other methods of treatment of cancer, including
CC leukaemias. The present sequence is the RBT1 protein
SQ Sequence 196 AA;

Query Match 100.0%; Score 1033; DB 4; Length 196;
Best Local Similarity 100.0%; Pred. No. 1.6e-92;
Matches 196; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 122 L-SHIEGLSQAPQ-----LADEGPPGRSIGGAAPSLGALDLLGPATGCLLDDGLEGLFE 175
Qy 153 DIDTSAVEKE---PARAPPEP-PHNLFCAPGSWE-----WNELDHIMEII 193
Db 176 DIDTSMYDNLWAPASEGLKPGPED---GPGKEEAPELDEAELDYLMDVL 222

Search completed: April 25, 2005, 12:29:28
Job time : 170 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 25, 2005, 12:23:30 ; Search time 41 Seconds
(without alignments)
459.963 Million cell updates/sec

Title: US-10-069-386A-2
Perfect score: 1033
Sequence: 1 MEGGLKRKHSDLSEEEERWE.....APGSWEWNELDHIMEIILGS 196

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79: *
1: pir1: *
2: pir2: *
3: pir3: *
4: pir4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	101.5	9.8	728	2	H59435	phosphoinositide-3
2	97	9.4	803	2	T40514	Chaperonin hsp78p
3	97	9.4	1201	2	G86441	unknown protein [i
4	96	9.3	817	2	S51342	verprolin - yeast
5	95.5	9.2	446	2	A42029	transcription fact
6	93.5	9.1	1734	2	A54602	microtubule-associ
7	93	9.0	433	1	FOLJH2	gag polyprotein -
8	91.5	8.9	1621	2	A82255	hypothetical prote
9	90.5	8.8	804	2	AG0565	probable membrane
10	88	8.5	651	2	T42644	hypothetical prote
11	88	8.5	716	2	T26998	hypothetical prote
12	88	8.5	1676	2	A56508	anucleate primary
13	86.5	8.4	536	2	A34596	transcription fact
14	86.5	8.4	1494	2	T14355	protein-tyrosine-p
15	86	8.3	475	2	B37761	ntrA protein - Thi
16	85.5	8.3	485	2	T37550	hypothetical coile
17	85.5	8.3	825	2	E75508	conserved hypothet
18	85.5	8.3	1544	2	E59431	phosphoinositide-b
19	85	8.2	715	2	D85087	hypothetical prote
20	85	8.2	1217	2	T00270	hypothetical prote
21	84.5	8.2	366	1	SAVLWD	large surface anti
22	84.5	8.2	477	2	T46304	hypothetical prote
23	84.5	8.2	723	2	B38749	3-phosphatidylinos
24	84.5	8.2	908	2	T16057	hypothetical prote
25	84.5	8.2	982	2	T43676	hunchback-related
26	84.5	8.2	1262	2	T13353	protein stn-B - fr
27	84	8.1	289	2	D70452	leucine-trNA ligas
28	84	8.1	954	2	I61714	co-repressor prote
29	84	8.1	1015	2	JC5062	phogrin precursor

RESULT 1

H59435

phosphoinositide-3-kinase regulatory beta chain [imported] - human

C;Species: Homo sapiens (man)

C;Date: 03-Jun-2002 #sequence_revision 03-Jun-2002 #text_change 09-Jul-2004

C;Accession: H59435; A59436

R;Volinia S; Patraccchini P; Otsu M; Hiles I; Gout I; Calzolari E; Bernardi F; Rooke L; We

Oncogene 7, 789-793, 1992

A;Title: Chromosomal localization of human p85 alpha, a subunit of phosphatidylinositol ;

A;Reference number: H59435

A;Accession: H59435

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-728 <VOL>

A;Cross-references: UNIPROT:O00459; GB:NP_005018; PID:g4826908; PIDN:NP_005018.1

R;Janssen, J.W.; Schleithoff, L.; Bartram, C.R.; Schulz, A.S.

Oncogene 16, 1767-1772, 1998

A;Title: An oncogenic fusion product of the phosphatidylinositol 3-kinase p85beta subunit

A;Reference number: A59436; MUID:98241181; PMID:9582025

A;Accession: A59436

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-728 <JAN>

A;Cross-references: GB:NP_005018; PID:g4826908; PIDN:NP_005018.1

Query Match

Best Local Similarity 9.8%; Score 101.5; DB 2; Length 728;

Matches 47; Conservativity 23.9%; Pred. No. 1;

Mismatches 25; Indels 70; Gaps 55; Gaps 9;

QY 20 EWSPA---GLQSYQQAL-----LRISLDKVQVSLGPRAPSLRRHVLIH 59

Db 164 QWDTAALADGIKSFLLALPAPLVTPEASAEARRALREAAGVGPALPPTLPLHRLTLR 223

QY 60 NTLQQLQAALRLAPALPPEPLFLGEEDFSLSATIGSILRELDTSMDGTPEPQNPTPL 119

Db 224 FLQLHLGRVARRAPA-----LGPVAVRALGATFGPLL-----LRAPPPSPSP-PFG 267

QY 120 GLQNEVPPQPD--PVFLEALSSRYLGDSDLDDFFLDIDTSAVEKEPARAPPEP----- 170

Db 268 GAPDGSESPDPFALLVEKLLQEHLEEQ-----EVAPPALPPKPPKAPPTVLANGG 320

QY 171 -PHNLFCAPGSWEWNEL 186

Db 321 SPPSL--QDAEWYWGDI 335

RESULT 2

T40514

Chaperonin hsp78p - fission yeast (Schizosaccharomyces pombe)

C;Species: Schizosaccharomyces pombe

C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004

C;Accession: T40514

Db 634 ASLLHDLVLPSSNLEKPPSPVAAAPPLPTFSAPSLPQQSVSTSI PPPV-APTLSVR-- 690

QY 143 GDSGLDDFFLDIDTSAVEKEPARAPPEPP 171

Db 691 -----TETESISKNPTKSPPPPP 708

RESULT 5
A42029
transcription factor E3 - mouse
C;Species: Mus musculus (house mouse)
C;Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C;Accession: A42029
R;Roman, C.; Matera, A.G.; Cooper, C.; Artandi, S.; Blain, S.; Ward, D.C.; Calame, K.
Mol. Cell. Biol. 12, 817-827, 1992
A;Title: mTFE3, an X-linked transcriptional activator containing basic helix-loop-helix
A;Reference number: A42029; MUID:92123207; PMID:1732746
A;Accession: A42029
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-446 <ROM>
A;Cross-references: UNIPROT:Q64092; GB:S76673; NID:G243439; PIDN:AAB21130.1; PID:G243440
A;Note: sequence extracted from NCBI backbone (NCBIN:76673, NCBI:76674)

Query Match 9.2%; Score 95.5; DB 2; Length 446;
Best Local Similarity 27.3%; Pred. No. 1.7;
Matches 48; Conservative 24; Mismatches 61; Indels 43; Gaps 8;

QY 12 LEEHEERWESPAGLQSYQQALLRISLDKVQSLGPRAPSLRRHVLHNTLQQLQAALRL 71

Db 272 LQKEQQR----SKDLESRQ-----SLEQANRSLQRLQEL-----ELQAIHG 311

QY 72 APAPALPPEPLFLGEEDFSATIGSILRELDTSMGTEPPQNPTPLGLQNEVPFPQDP 131

Db 312 LP----VPPNPGLLSLTSSVSDSLKP--EQLDIEEGREPTTFHVSGGPAQNAPPQQPPA 366

QY 132 VFLEAL-----SSRYLGDGDDFFLDIDTSAVEKE-----PARAPEP 170

Db 367 PPSDALDLHFPSPDHLGDLG-DPFHLGLEDI LMEEGMVGGSLGGALSPLRAASDP 421

RESULT 6
A54602
microtubule-associated serine/threonine protein kinase MAST205 - mouse
C;Species: Mus musculus (house mouse)
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004
C;Accession: A54602
R;Walden, P.D.; Cowan, N.J.
Mol. Cell. Biol. 13, 7625-7635, 1993
A;Title: A novel 205-kilodalton testis-specific serine/threonine protein kinase associat
A;Reference number: A54602; MUID:94067123; PMID:8246979
A;Accession: A54602
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1734 <WAL>
A;Cross-references: UNIPROT:Q60592; GB:U02313; NID:G406057; PIDN:AAC04312.1; PID:G406058
C;Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo
C;Keywords: ATP
F;451-726/Domain: protein kinase homology <KIN>
F;459-467/Region: protein kinase ATP-binding motif

Query Match 9.1%; Score 93.5; DB 2; Length 1734;
Best Local Similarity 26.3%; Pred. No. 15;
Matches 51; Conservative 23; Mismatches 71; Indels 49; Gaps 8;

QY 25 GLQSYQQAL-----LRISLDKVQSLGPRAPSLRRHVLHNTLQQLQAALRLAPAPALPP 79

Db 1308 GSQSFPTKHLSPPLGRQLSRPKSAEPPRSPLLR-----VQSAEKLAAALAAAE 1357

QY 80 EPLFLGEEDFSLSATIGSILRELDTSMGTEPPQNPTPLGLQNEVP---PQDPVFLEA 136

Db 1358 KKL-APSRKHSLDLPHGELKKEL-----TPREASPLEVVGTRSVLSGKPLPGKGVLPQ 1410

QY 137 LSSRYLG-----DSGLDDFFLDIDTSAVEKEPARAP-PEPPHN 173

Db 1411 APSRALGTLRQDRAERRESLQKQEAIREVDSSEDDTDEPENSOATQERLSPHPPEASHN 1470

QY 174 LFCAPGSWEWNELD 187

Db 1471 LL-PKSGEGTEED 1483

RESULT 7
FOLJH2
gag polyprotein - human T-cell lymphotropic virus type 2
N;Alternate names: core polyprotein
N;Contains: core protein p12; core protein p15; core protein p24
C;Species: human T-cell lymphotropic virus type 2, HTLV-2
A;Note: host Homo sapiens (man)
C;Date: 28-Feb-1986 #sequence_revision 28-Feb-1986 #text_change 09-Jul-2004
C;Accession: A03944
R;Shimotohno, K.; Takahashi, Y.; Shimizu, N.; Gojobori, T.; Golde, D.W.; Chen, I.S.Y.; Mi
Proc. Natl. Acad. Sci. U.S.A. 82, 3101-3105, 1985
A;Title: Complete nucleotide sequence of an infectious clone of human T-cell leukemia vi
A;Reference number: A94042; MUID:85216449; PMID:2582407
A;Accession: A03944
A;Molecule type: DNA
A;Residues: 1-433 <SHI>
A;Cross-references: UNIPROT:P03346; GB:M10060; NID:G329559; PIDN:AAB59884.1; PID:G329562
A;Note: the authors translated the codon TAC for residues 197 and 249 as Thr
C;Genetics:
A;Gene: gag
C;Superfamily: mammalian retrovirus gag polyprotein II
C;Keywords: core protein; polyprotein
F;1-136/Product: core protein p15 #status predicted <P15>
F;137-214/Product: core protein p24 #status predicted <P24>
F;215-433/Product: core protein p12 #status predicted <P12>

Query Match 9.0%; Score 93; DB 1; Length 433;
Best Local Similarity 23.1%; Pred. No. 2.7;
Matches 40; Conservative 24; Mismatches 67; Indels 42; Gaps 7;

QY 47 PRAPSLRRHVLHNTLQQLQAALRLAPAPA-----LPPEPLFLGEEDFSLSA 93

Db 13 PKAP---RGLSTHHLNLFQAAYRLQRPSPDFDFQRLRFLKALKTPILWLNPIDYSLLA 69

QY 94 TI-----GSILRELDTSMGTEPPQNPTPLG---LQNEVPPQDPVFLEA-LSSRYLG 143

Db 70 SLIPKGYPGRVVEIINILVKNQVSPSAPAAPVPTPICPTTTTPPPPPSPEAHVPPPY-- 127

QY 144 DSGLDDFFLDIDTSAVEKEPARAPPEPPHNLFCAPGSWEWNELDHIMEIILGS 196

Db 128 -----VEPTTQCFFPILHPPGAP----SAHRPWQMKDLQAIKQEVSSS 166

RESULT 8
A82255
hypothetical protein VC0998 [imported] - Vibrio cholerae (strain N16961 serogroup O1)
C;Species: Vibrio cholerae
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C;Accession: A82255
R;Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P.
l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A;Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A;Reference number: A82035; MUID:20406833; PMID:10952301
A;Accession: A82255
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1621 <HEI>
A;Cross-references: UNIPROT:Q9KTA5; GB:AE004181; GB:AE003852; NID:G9655454; PIDN:AAF9415;
A;Experimental source: serogroup O1; strain N16961; biotype El Tor
C;Genetics:
A;Gene: VC0998

A;Molecule type: DNA
A;Residues: 1-1676 <FIS>
A;Cross-references: UNIPROT:Q00083; GB:X82289; NID:G683499; PID:G683500
C;Genetics:
A;Gene: apsa
A;Introns: 149/3

Query Match 8.5%; Score 88; DB 2; Length 1676;
Best Local Similarity 20.5%; Pred. No. 42;
Matches 32; Conservative 19; Mismatches 51; Indels 54; Gaps 6;

QY 36 ISLDKQVRSGLGRAPSLRRHVLHNTLQQLQAALRLAPAPALPPEPLFLGEEDEFSLSATI 95
Db 837 VSFEETE-PVAPSPPELRTAFFVGSTTEPVAAAPVPVPEVALSP-----I 880

QY 96 GSILRELDTSMDGTEPPQNVPVPLGLQNEVPPQDPVFL----- 134
Db 881 SS-----QTQTPTFPVIP-----APPEPEPIYVPEMAFSQILVEDTLPILAKLP 924

QY 135 EALSSRYLGDSGLDDFFLDIDTSAVEKEPARAPPEP 170
Db 925 EPAPERVFAEQGTSTDIADVSVSAISSEQTE-PVEP 959

RESULT 13
A34596
transcription factor E3 - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 06-Jul-1990 #sequence_revision 09-Oct-1992 #text_change 09-Jul-2004
C;Accession: A34596; S10379
R;Beckmann, H.; Su, L.K.; Kadesch, T.
Genes Dev. 4, 167-179, 1990
A;Title: TFE3: a helix-loop-helix protein that activates transcription through the immunoglobulin enhancer
A;Reference number: A34596; MUID:90249724; PMID:2338243
A;Accession: A34596
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-536 <BEC>
A;Cross-references: UNIPROT:P19532; EMBL:X51330; NID:G37061; PIDN:CAA35714.1; PID:gl3353
A;Note: the authors translated the codon ACC for residue 433 as Ser, and GAG for residue 534
C;Genetics:
A;Gene: GDB:TFE3
A;Cross-references: GDB:125870; OMIM:314310
A;Map position: Xp11.23-Xp11.22
C;Keywords: DNA binding; transcription factor

Query Match 8.4%; Score 86.5; DB 2; Length 536;
Best Local Similarity 25.4%; Pred. No. 13;
Matches 47; Conservative 24; Mismatches 55; Indels 59; Gaps 10;

QY 12 LEEEEEERWESPAGLQSYQALLRISLDKQVRSGLGRAPSLRRHVLHNTLQQLQAALRL 71
Db 192 LQKEQQR----SKDLESQR----SLEQANRSLQRIQEL-----ELQAQIHG 231

QY 72 APAPALPPEPLFLGEEDEFSLSATIGSIILRE-LDTSMDGTE-----PPQNVPVPLG 120
Db 232 LPVPVGT-----GLLSLATSTSDSLKPEQLDIEEGRPGARTFHVGGGPAQN----- 279

QY 121 LQNEVPPQP--DPVFLEALSSRYLGDSGLDDFFLDIDTSAVEKE-----PAR 165
Db 280 APHQQPPAPPSSDALLDLHPFSDHLGDLG-DPFFHLGLELDILMEEEGVVGLSGGALSPLR 338

QY 166 APPEP 170
Db 339 AASDP 343

RESULT 14
T14355
protein-tyrosine-phosphatase (EC 3.1.3.48) TD14 - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
C;Accession: T14355

R;Cao, L.; Zhang, L.; Ruiz-Lozano, P.; Yang, Q.; Chien, K.R.; Graham, R.M.; Zhou, M.
J. Biol. Chem. 273, 21077-21083, 1998
A;Title: A novel putative protein-tyrosine phosphatase contains a BRO1-like domain and is related to the SHPTP family
A;Reference number: Z18004; MUID:98361981; PMID:9694860
A;Accession: T14355
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-1494 <CAO>
A;Cross-references: UNIPROT:O88902; EMBL:AF077000; NID:G3598973; PID:G3598974; PIDN:AAC6;
A;Experimental source: brain
C;Genetics:
A;Gene: PTP-TD14
C;Function:
A;Description: may be involved in regulating Ha-ras-dependent cell growth
C;Keywords: phosphoric monoester hydrolase

Query Match 8.4%; Score 86.5; DB 2; Length 1494;
Best Local Similarity 29.1%; Pred. No. 48;
Matches 37; Conservative 19; Mismatches 40; Indels 31; Gaps 9;

QY 62 LQQLQAALRLAPAPALPPEPLFLGEEDEFSLSATIGSIILRE-----LDTSMDGTEPPQNP 115
Db 625 LMQPRAAVPMAPGVLVPAPVYTSE-----LGLVPRSSPQHGVSSPYAGVGPQP-P 675

QY 116 VTPLGLQNEVPPQ---PDPVFLEALSSRYLGDSGLDDFFLDIDT-SAVEKEPARAPPEPP 171
Db 676 I--VGLPSAPPQPSGPE---LAMDV-R-PATTVDSVQAPISSHMLALRPGAPAPQP- 727

QY 172 HNLFCAP 178
Db 728 ----CFP 730

RESULT 15
B37761
ntrA protein - Thiobacillus ferrooxidans
C;Species: Thiobacillus ferrooxidans
C;Date: 31-May-1991 #sequence_revision 31-May-1991 #text_change 09-Jul-2004
C;Accession: B37761
R;Berger, D.K.; Woods, D.R.; Rawlings, D.E.
J. Bacteriol. 172, 4399-4406, 1990
A;Title: Complementation of Escherichia coli sigma(54) (ntrA)-dependent formate hydrogen lyase
A;Reference number: A37761; MUID:90330545; PMID:2198257
A;Accession: B37761
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-475 <BER>
A;Cross-references: UNIPROT:P24695; GB:M58480; GB:M33831; NID:gl54642; PIDN:AAA27379.1; PID:gl3353
C;Superfamily: Pseudomonas transcription initiation factor sigma
C;Keywords: DNA binding; transcription regulation

Query Match 8.3%; Score 86; DB 2; Length 475;
Best Local Similarity 21.7%; Pred. No. 12;
Matches 40; Conservative 31; Mismatches 75; Indels 38; Gaps 6;

QY 1 MEGGLKRRKHSDLSEEEERWESPAGLQSYQALLRISLDKQVQ-----SLGPRA--PSLR 53
Db 307 MAGGKDAAHKYIQDQLNEARWFIKSLQSRQDTILKVARAIVERQKOFFANGPESMRPMVL 366

QY 54 RHVLIHNTLQQLQAALRLAPAPALPPEPLFLGEEDEFSLSATIG-----SILRELDT 104
Db 367 RH--IADAVEHVESTVSRVTNQYMITPRGLYEPKVFYFFSSHVGTGDSGGSASATAIRALLI 424

QY 105 SMDGTEPPQNVPVPLGLQNEVPPQDPVFLEALSSRYLGDSGLDDFFLDIDTSAVEKEBP 164
Db 425 KMTQAEADAQHPLS-----DAEIRVLADQGIQ---IARRTVAKYVREAA 464

QY 165 RAPP 168
Db 465 NVPP 468

Search completed: April 25, 2005, 12:33:15

Job time : 44 secs



GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 25, 2005, 12:32:36 ; Search time 134 Seconds
(without alignments)
486.769 Million cell updates/sec

Title: US-10-069-386A-2
Perfect score: 1033
Sequence: 1 MEGGLKRXKHSDLSEEEERWE.....APGSWEWNELDHIMEILGS 196

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1424015 seqs, 332791073 residues

Total number of hits satisfying chosen parameters: 1424015

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	791	76.6	157	14	US-10-106-698-5516	Sequence 5516, Ap
2	645.5	62.5	142	14	US-10-029-386-33709	Sequence 33709, A
3	402	38.9	111	14	US-10-029-386-33708	Sequence 33708, A
4	216	20.9	236	11	US-09-978-360A-568	Sequence 568, App
5	213	20.6	236	9	US-09-731-872-412	Sequence 412, App
6	213	20.6	236	10	US-09-876-997-412	Sequence 412, App
7	213	20.6	278	15	US-10-296-115-1065	Sequence 1065, Ap
8	205	19.8	222	14	US-10-029-386-34001	Sequence 34001, A
9	136	13.2	237	13	US-10-076-069-2	Sequence 2, Appli
10	125	12.1	314	15	US-10-170-385-31	Sequence 31, Appl
11	125	12.1	314	16	US-10-755-889-627	Sequence 627, App
12	106	10.3	254	9	US-09-925-302-596	Sequence 596, App
13	106	10.3	254	10	US-09-925-302-596	Sequence 596, App

14	106	10.3	254	14	US-10-106-698-4456	Sequence 4456, Ap
15	105	10.2	241	13	US-10-076-069-4	Sequence 4, Appli
16	104.5	10.1	740	15	US-10-276-774-2083	Sequence 2083, Ap
17	98	9.5	562	15	US-10-104-047-3214	Sequence 3214, Ap
18	97	9.4	803	15	US-10-369-493-2536	Sequence 2536, Ap
19	96	9.3	297	15	US-10-425-114-64002	Sequence 64002, A
20	96	9.3	817	15	US-10-369-493-1813	Sequence 1813, Ap
21	95.5	9.2	1228	16	US-10-437-963-188300	Sequence 188300,
22	95	9.2	713	16	US-10-437-963-137248	Sequence 137248,
23	94	9.1	433	8	US-08-259-451-3	Sequence 3, Appli
24	94	9.1	433	14	US-10-224-999A-3472	Sequence 3472, Ap
25	93.5	9.1	256	15	US-10-421-138A-312	Sequence 312, App
26	93.5	9.1	256	15	US-10-374-780A-1247	Sequence 1247, Ap
27	93.5	9.1	663	16	US-10-755-889-480	Sequence 480, App
28	93.5	9.1	1734	9	US-09-862-027-81	Sequence 81, Appl
29	93.5	9.1	1734	15	US-10-042-865-82	Sequence 82, Appl
30	93	9.0	157	16	US-10-767-701-54984	Sequence 54984, A
31	92.5	9.0	113	13	US-10-076-069-6	Sequence 6, Appli
32	92	8.9	1103	16	US-10-437-963-166325	Sequence 166325,
33	91.5	8.9	714	15	US-10-425-114-62753	Sequence 62753, A
34	91	8.8	316	15	US-10-425-114-56782	Sequence 56782, A
35	90.5	8.8	206	16	US-10-437-963-187376	Sequence 187376,
36	90.5	8.8	370	14	US-10-192-381-16	Sequence 16, Appl
37	90.5	8.8	824	14	US-10-226-844-1	Sequence 1, Appli
38	90.5	8.8	824	14	US-10-210-951-58	Sequence 58, Appl
39	90.5	8.8	824	14	US-10-211-884-58	Sequence 58, Appl
40	90.5	8.8	824	15	US-10-211-858-58	Sequence 58, Appl
41	90.5	8.8	824	16	US-10-789-378-4	Sequence 4, Appli
42	90.5	8.8	824	17	US-10-741-600-1079	Sequence 1079, Ap
43	90.5	8.8	947	9	US-09-871-889-1	Sequence 1, Appli
44	90.5	8.8	947	10	US-09-981-397A-18	Sequence 18, Appl
45	90.5	8.8	947	13	US-10-087-192-888	Sequence 888, App

ALIGNMENTS

RESULT 1

US-10-106-698-5516
; Sequence 5516, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptides
; FILE REFERENCE: PA005P1
; CURRENT APPLICATION NUMBER: US/10/106,698
; PRIOR FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 5516
; LENGTH: 157
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (55)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: MISC_FEATURE
; LOCATION: (132)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-106-698-5516

Query Match 76.6%; Score 791; DB 14; Length 157;
Best Local Similarity 98.0%; Pred. No. 9.2e-64;
Matches 150; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 44 SLGPRAPSLRRHVLIHNTLQQLQAALRLAPALPPEPLFLGEEDFSLSATIGSILRELD 103

```
Db      :||||| 5 ALGPRAPSLRRHVLHNTLQQLQAALRLAPAPALPPEPLFLGEEDFSL SAXIGSILRELD 64
QY      :||||| 104 TSMGTEPPQNPVTPLGLQNEVPPQDPDPVFLEALSSRYLGDSDLDDFFLDIDTSAVEKEP 163
Db      :||||| 65 TSMGTEPPQNPVTPLGLQNEVPPQDPDPVFLEALSSRYLGDSDLDDFFLDIDTSAVEKEP 124
QY      :||||| 164 ARAPPEPPHNLFCAPGSWEWNELDHIMEIILGS 196
Db      :||||| 125 ARAPPEPXHNLFCAPGSWEWNELDHIMEIILGS 157

RESULT 2
US-10-029-386-33709
; Sequence 33709, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 33709
; LENGTH: 142
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC010271.5
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.64
US-10-029-386-33709
Query Match      62.5%; Score 645.5; DB 14; Length 142;
Best Local Similarity 96.8%; Pred. No. 1.3e-50;
Matches 122; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY      :||||| 71 LAPAPALPPEPLFLGEEDFSL SATIGSILRELDTSMGTEPPQNPVTPLGLQNEVPPQPD 130
Db      :||||| 18 LPPLPC-PPEPLFLGEEDFSL SATIGSILRELDTSMGTEPPQNPVTPLGLQNEVPPQPD 76
QY      :||||| 131 PVFLEALSSRYLGDSDLDDFFLDIDTSAVEKEPARAPPEPPHNLFCAPGSWEWNELDHIM 190
Db      :||||| 77 PVFLEALSSRYLGDSDLDDFFLDIDTSAVEKEPARAPPEPPHNLFCAPGSWEWNELDHIM 136
QY      :||||| 191 EIILGS 196
Db      :||||| 137 EIILGS 142

RESULT 3
US-10-029-386-33708
; Sequence 33708, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 33708
; LENGTH: 111
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
```

```
; OTHER INFORMATION: MAP TO AC010271.5
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.64
US-10-029-386-33708
Query Match      38.9%; Score 402; DB 14; Length 111;
Best Local Similarity 89.9%; Pred. No. 1.1e-28;
Matches 80; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY      :||||| 1 MEGGLKRXKHSDLSEEEERWESWSPAGLQSYQQALLRISLDKVQVSLGPRAPSLRRHVLHN 60
Db      :||||| 3 MVGGLKRXKHSDLSEEEERWESWSPAGLQSYQQALLRISLDKVQVSLGPRAPSLRRHVLHN 62
QY      :||||| 61 TLQQLQAALRLAPAPALPPEPLFLGEEDF 89
Db      :||||| 63 TLQQLQAALRLAPAPALPPRAPLPGRGGF 91

RESULT 4
US-09-978-360A-568
; Sequence 568, Application US/09978360A
; Publication No. US20040110939A1
; GENERAL INFORMATION:
; APPLICANT: Edwards, Jean-Baptiste Dumas Milne
; APPLICANT: Duclert, Aymeric
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Jobert, Severin
; APPLICANT: Clusel, Catherine
; TITLE OF INVENTION: Complementary DNA's Encoding Proteins with Signal Peptides
; FILE REFERENCE: 56.US4.CIP
; CURRENT APPLICATION NUMBER: US/09/978,360A
; CURRENT FILING DATE: 2001-10-15
; PRIOR APPLICATION NUMBER: US 60/066,677
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: US 60/069,957
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: US 60/074,121
; PRIOR FILING DATE: 1998-02-09
; PRIOR APPLICATION NUMBER: US 60/081,563
; PRIOR FILING DATE: 1998-04-13
; PRIOR APPLICATION NUMBER: US 60/096,116
; PRIOR FILING DATE: 1998-08-10
; PRIOR APPLICATION NUMBER: US 60/099,273
; PRIOR FILING DATE: -09-04
; PRIOR APPLICATION NUMBER: US 09/191,997
; PRIOR FILING DATE: 1998-11-13
; PRIOR APPLICATION NUMBER: US 09/215,435
; PRIOR FILING DATE: 1998-12-17
; PRIOR APPLICATION NUMBER: PCT/IB98/02122
; PRIOR FILING DATE: 1998-12-17
; PRIOR APPLICATION NUMBER: US 09/247,155
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 810
; SOFTWARE: Patent.pm
; SEQ ID NO 568
; LENGTH: 236
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -31...-1
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (28, 30, 40, 67, 86, 117, 120)
; OTHER INFORMATION: unknown
US-09-978-360A-568
Query Match      20.9%; Score 216; DB 11; Length 236;
Best Local Similarity 32.6%; Pred. No. 2.1e-11;
Matches 76; Conservative 36; Mismatches 69; Indels 52; Gaps 13;

QY      :||||| 4 GLKRXKHSDLSEEE---ERWESWSPAGLQSYQQA-----LLRISLDKVQVSLGPRAPS 51
; :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
```


Db 5 GLKRKREEEKEPLAVDSW-WLDPGHAAVAQAAPPVASSSLFDLSVLKHLHSLQXSXP 63

QY 52 LRRHVLHNTLQQLQAALRLAPALPPEPL-----FLGEEDFSLSATIGSILRE 101

Db 64 LRHLVLVXNTLRRRIQAS--MAPAAALPPVPTPPAAPXVADNLLASSDAALSASMAXLLED 121

QY 102 LDTSMGTEPPQNPTPLGLQNEVPQPD-----PVPFLEAL-----SSRYLGDSGLDDFFL 152

Db 122 L-SHIEGLSQAQP-----LADEGPPGRSIGGXPPXIGALDLLGPATGCLLDNGLEGLFE 175

QY 153 DIDTSAVEKE---PARAPPEP-PHNLFCAPGSWE-----WNELDHIMEIILGS 196

Db 176 DIDTSMYDNLWAPASEGLKPGPED---GPGKEEAPDELDEAELDYLMMDVLVGT 225

RESULT 5

US-09-731-872-412

; Sequence 412, Application US/09731872

; Patent No. US20020102604A1

; GENERAL INFORMATION:

; APPLICANT: Dumas Milne Edwards, Jean Baptiste

; APPLICANT: Bougueleret, Lydie

; APPLICANT: Jobert, Severin

; TITLE OF INVENTION: FULL-LENGTH HUMAN cDNAs ENCODING POTENTIALLY SECRETED PROTEINS

; FILE REFERENCE: 78.US3.REG

; CURRENT APPLICATION NUMBER: US/09/731,872

; CURRENT FILING DATE: 2000-12-07

; PRIOR APPLICATION NUMBER: US 60/169,629

; PRIOR FILING DATE: 1999-12-08

; PRIOR APPLICATION NUMBER: US 60/187,470

; PRIOR FILING DATE: 2000-03-06

; NUMBER OF SEQ ID NOS: 482

; SOFTWARE: Patent.pm

; SEQ ID NO 412

; LENGTH: 236

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-731-872-412

Query Match 20.6%; Score 213; DB 9; Length 236;

Best Local Similarity 32.6%; Pred. No. 3.9e-11;

Matches 76; Conservative 36; Mismatches 69; Indels 52; Gaps 13;

QY 4 GLKRKHSDELEEE---ERWESPAGLQSYQQA-----LLRISLDKVQVRSIGPRAPS 51

Db 5 GLKRKREEEKEPLAVDSW-WLDPGHTAVAQAAPPVASSSLFDLSVLKHLHSLQQSEPD 63

QY 52 LRRHVLHNTLQQLQAALRLAPALPPEPL-----FLGEEDFSLSATIGSILRE 101

Db 64 LRHLVLVNTLRRRIQAS--MAPAAALPPVPSPPAAPSVDNLLASSDAALSASMASLLED 121

QY 102 LDTSMGTEPPQNPTPLGLQNEVP-----QDPVFLEAL---SSRYLGDSGLDDFFL 152

Db 122 L-SHIEGLSQAQP-----LADEGPPGRSIGGAAPSLGALDLLGPATGCLLDGLEGLE 175

QY 153 DIDTSAVEKE---PARAPPEP-PHNLFCAPGSWE-----WNELDHIMEIILGS 196

Db 176 DIDTSMYDNLWAPASEGLKPGPED---GPGKEEAPDELDEAELDYLMMDVLVGT 225

RESULT 6

US-09-876-997-412

; Sequence 412, Application US/09876997

; Publication No. US20030152921A1

; GENERAL INFORMATION:

; APPLICANT: Dumas Milne Edwards, Jean Baptiste

; APPLICANT: Bougueleret, Lydie

; APPLICANT: Jobert, Severin

; TITLE OF INVENTION: FULL-LENGTH HUMAN cDNAs ENCODING POTENTIALLY SECRETED PROTEINS

; FILE REFERENCE: 78.US4.CIP

; CURRENT APPLICATION NUMBER: US/09/876,997

; CURRENT FILING DATE: 2001-06-08

; PRIOR APPLICATION NUMBER: US 09/731,872

; PRIOR FILING DATE: 2000-12-07

; PRIOR APPLICATION NUMBER: US 60/187,470

; PRIOR FILING DATE: 2000-03-06

; PRIOR APPLICATION NUMBER: US 60/169,629

; PRIOR FILING DATE: 1999-12-08

; NUMBER OF SEQ ID NOS: 482

; SOFTWARE: Patent.pm

; SEQ ID NO 412

; LENGTH: 236

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-876-997-412

Query Match 20.6%; Score 213; DB 10; Length 236;

Best Local Similarity 32.6%; Pred. No. 3.9e-11;

Matches 76; Conservative 36; Mismatches 69; Indels 52; Gaps 13;

QY 4 GLKRKHSDELEEE---ERWESPAGLQSYQQA-----LLRISLDKVQVRSIGPRAPS 51

Db 5 GLKRKREEEKEPLAVDSW-WLDPGHTAVAQAAPPVASSSLFDLSVLKHLHSLQQSEPD 63

QY 52 LRRHVLHNTLQQLQAALRLAPALPPEPL-----FLGEEDFSLSATIGSILRE 101

Db 64 LRHLVLVNTLRRRIQAS--MAPAAALPPVPSPPAAPSVDNLLASSDAALSASMASLLED 121

QY 102 LDTSMGTEPPQNPTPLGLQNEVP-----QDPVFLEAL---SSRYLGDSGLDDFFL 152

Db 122 L-SHIEGLSQAQP-----LADEGPPGRSIGGAAPSLGALDLLGPATGCLLDGLEGLE 175

QY 153 DIDTSAVEKE---PARAPPEP-PHNLFCAPGSWE-----WNELDHIMEIILGS 196

Db 176 DIDTSMYDNLWAPASEGLKPGPED---GPGKEEAPDELDEAELDYLMMDVLVGT 225

RESULT 7

US-10-296-115-1065

; Sequence 1065, Application US/10296115

; Publication No. US20040053248A1

; GENERAL INFORMATION:

; APPLICANT: Hyseq Inc

; TITLE OF INVENTION: No. US20040053248A1el Nucleic Acids and Polypeptides

; FILE REFERENCE: 784PCT

; CURRENT APPLICATION NUMBER: US/10/296,115

; CURRENT FILING DATE: 2002-11-18

; PRIOR APPLICATION NUMBER: US09/488,725

; PRIOR FILING DATE: 2000-01-21

; PRIOR APPLICATION NUMBER: US09/552,317

; PRIOR FILING DATE: 2000-04-25

; NUMBER OF SEQ ID NOS: 1478

; SEQ ID NO 1065

; LENGTH: 278

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-296-115-1065

Query Match 20.6%; Score 213; DB 15; Length 278;

Best Local Similarity 32.6%; Pred. No. 4.8e-11;

Matches 76; Conservative 36; Mismatches 69; Indels 52; Gaps 13;

QY 4 GLKRKHSDELEEE---ERWESPAGLQSYQQA-----LLRISLDKVQVRSIGPRAPS 51

Db 47 GLKRKREEEKEPLAVDSW-WLDPGHAAVAQAAPPVASSSLFDLSVLKHLHSLQQSEPD 105

QY 52 LRRHVLHNTLQQLQAALRLAPALPPEPL-----FLGEEDFSLSATIGSILRE 101

Db 106 LRHLVLVNTLRRRIQAS--MAPAAALPPVPSPPAAPSVDNLLASSDAALSASMASLLED 163

QY 102 LDTSMGTEPPQNPTPLGLQNEVP-----QDPVFLEAL---SSRYLGDSGLDDFFL 152

Db 164 L-SHIEGLSQAQP-----LADEGPPGRSIGGAAPSLGALDLLGPATGCLLDGLEGLE 217

QY 153 DIDTSAVEKE---PARAPPEP-PHNLFCAPGSWE-----WNELDHIMEIILGS 196

QY 58 IHNTLQQLQAAL-----RLAP--APALPPEPLFLGEEEDFSLSATIGSILRELDTSMDG 108
| | | | | : | | | | | : | | | | | : | | | | | :
Db 75 IANTVRQIQEEMTQDGTWRTVAPQAAERAPXDRLVSTE-----ILCRAAWGQEG 123

QY 109 TEP-----PQNPVT-----PLGLQNEVPPQDP-----VFLEALS 138
| | | | | : | | | | | : | | | | | : | | | | | :
Db 124 AHPAPGLGDGHTQGPVSDLCPV TSAQAPRHLQSSAWEMDGPRENRGSFHKSLDQIFETLE 183

QY 139 SRYLGDSDLDDFFLDIDTSAVEKEP-----AR-----APPEPHNLFCA PGS 180
:
Db 184 TK--NPSCMEELFSDVSPYYDLDTVLTGMMGGARPGPCEGLEGLAPATPGPSSSCKS-- 239

QY 181 WEWNELDHIMEIIL 194
: | | | | : | | | :
Db 240 -DLGELDHVVEILV 252

RESULT 14
US-10-106-698-4456
; Sequence 4456, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide
; FILE REFERENCE: PA005P1
; CURRENT APPLICATION NUMBER: US/10/106,698
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 4456
; LENGTH: 254
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (105)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-106-698-4456

Query Match 10.3%; Score 106; DB 14; Length 254;
Best Local Similarity 25.6%; Pred. No. 0.21;
Matches 65; Conservative 32; Mismatches 75; Indels 82; Gaps 13;

QY 4 GLKRKHSDL EEEERWESPAGLQ---SY---QQALLRISLDKVORS LGPRAPSLRRHVL 57
| | | | | : | | | | | : | | | | | : | | | | | :
Db 18 GLKRKCVGHEEDV---EGALAGLKT VSSYSLQ RQSLDMSLVKLQ LCHMLVEPNLCRSVL 74

QY 58 IHNTLQQLQAAL-----RLAP--APALPPEPLFLGEEEDFSLSATIGSILRELDTSMDG 108
| | | | | : | | | | | : | | | | | : | | | | | :
Db 75 IANTVRQIQEEMTQDGTWRTVAPQAAERAPXDRLVSTE-----ILCRAAWGQEG 123

QY 109 TEP-----PQNPVT-----PLGLQNEVPPQDP-----VFLEALS 138
| | | | | : | | | | | : | | | | | : | | | | | :
Db 124 AHPAPGLGDGHTQGPVSDLCPV TSAQAPRHLQSSAWEMDGPRENRGSFHKSLDQIFETLE 183

QY 139 SRYLGDSDLDDFFLDIDTSAVEKEP-----AR-----APPEPHNLFCA PGS 180
:
Db 184 TK--NPSCMEELFSDVSPYYDLDTVLTGMMGGARPGPCEGLEGLAPATPGPSSSCKS-- 239

QY 181 WEWNELDHIMEIIL 194
: | | | | : | | | :
Db 240 -DLGELDHVVEILV 252

RESULT 15
US-10-076-069-4
; Sequence 4, Application US/10076069

; Publication No. US20020177214A1
; GENERAL INFORMATION:
; APPLICANT: JURECIC, ROLAND
; APPLICANT: NACHTMAN, RONALD
; TITLE OF INVENTION: HEPP, A NOVEL GENE WITH A ROLE IN HEMATOPOIETIC AND NEURAL DEVELOPMENT
; FILE REFERENCE: 39532-176599
; CURRENT APPLICATION NUMBER: US/10/076,069
; CURRENT FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: US 60/268,923
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 241
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-076-069-4

Query Match 10.2%; Score 105; DB 13; Length 241;
Best Local Similarity 25.9%; Pred. No. 0.25;
Matches 65; Conservative 31; Mismatches 79; Indels 76; Gaps 13;

QY 4 GLKRKHSDL EEEERWESPAGLQ---SY---QQALLRISLDKVORS LGPRAPSLRRHVL 57
| | | | | : | | | | | : | | | | | : | | | | | :
Db 5 GLKRKCVGHEEDV---EGALAGLKT VSSYSLQ RQSLDMSLVKLQ LCHMLVEPNLCRSVL 61

QY 58 IHNTLQQLQAAL-----RLAPAPALPPEPLFLGEEEDFSLSATIGSILRELDTSMDGTEP 111
| | | | | : | | | | | : | | | | | : | | | | | :
Db 62 IANTVRQIQEEMTQDGTWRTVAPQAAERAPL-----DRLVST---EILCRAAWGQEG AHP 113

QY 112 -----PQNPVT-----PLGLQNEVPPQDP-----VFLEALSSRY 141
| | | | | : | | | | | : | | | | | : | | | | | :
Db 114 APGLGDGHTQGPVSDLCPV TSAQAPRHLQSSAWEMDGPRENRGSFHKSLDQIFETLETK- 172

QY 142 LGDSGLDDFFLDIDTSAVEKEP-----AR-----APPEPHNLFCA PGSWEW 183
:
Db 173 -NPSCMEELFSDVSPYYDLDTVLTGMMGGARPGPCEGLEGLAPATPGPSSSCKS---DL 228

QY 184 NELDHIMEIIL 194
| | | | | : | | | :
Db 229 GELDHVVEILV 239

Search completed: April 25, 2005, 12:45:21
Job time : 135 secs


```
; APPLICANT: Dougherty, Brian A
; APPLICANT: Goldberg, Steven L
; APPLICANT: Hofle, Gerhard
; APPLICANT: Mueller, Joachim
; APPLICANT: Reichenbach, Hans
; TITLE OF INVENTION: DNA sequences for enzymatic synthesis of polyketide or
; TITLE OF INVENTION: heteropolyketide compounds
; FILE REFERENCE: PCT/US 99/23535
; CURRENT APPLICATION NUMBER: US/09/413,814
; CURRENT FILING DATE: 1999-10-07
; EARLIER APPLICATION NUMBER: DE 198 46 493.2
; EARLIER FILING DATE: 1998-10-09
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 28
; LENGTH: 2618
; TYPE: PRT
; ORGANISM: Sorangium cellulosum
US-09-413-814-28

Query Match          9.0%; Score 93; DB 3; Length 2618;
Best Local Similarity 25.5%; Pred. No. 9.2;
Matches 55; Conservative 18; Mismatches .79; Indels 64; Gaps 11;

QY      31 QALLR---ISLD-----KVQRSIGPRAP--SLRRHVLIHNTLQQLOA 67
       ||||| | ||
DB       974 QALLRREAIGLDEPFPOAGNSFGLIRLHAKLESAPKSPFITDLFQHTSIRSQAEMLSG 1033
               : : : | | | : | : : | : : |
QY      68 ALRLAP-APALPPEPLFLGEEDFSLSATIGSILRELDTSMDGTTEPPQ---NPVTPLGLQN 123
       : ||| | : | : : | : | : | : | : | : | : | : |
DB     1034 SSVEAPLAGAVPQPAAAQAQVASSAAKSPGERGAATSSGLTAQPPQPHFRPIAVIGLAG 1093
               : ||| | : | : : | : | : | : | : | : | : | : |
QY     124 EVPPQPD-PVFLEAL-----SSRYLGDSGLD-----DPFLD 153
       - ||| | ||| | | | | | | | | | | | | | | | |
DB     1094 RFPAAPDLDAFLELLTEGRCGIRFFFSQAELRDEGLDANRIACHNVYPKGFDRADHF-D 1152
               - ||| | | | | | | | | | | | | | | | |
QY     154 IDTSAVEKEPAR-APPEPPHNLFCAPGSWENWELDH 188
       | : | : | : | : | : | : | : | : | : | : | : |
DB     1153 ADFFGIPPRDAEITDQIIRLLLECC-----WNALEH 1183
```

RESULT 5
PCT-US95-04910-13
; Sequence 13, Application PC/TUS9504910
; GENERAL INFORMATION:
; APPLICANT: The Government of the United
; APPLICANT: States of America as represented
; APPLICANT: by the Secretary, Department of
; APPLICANT: Health and Human Services
; TITLE OF INVENTION: ISOLATION AND
; TITLE OF INVENTION: CHARACTERIZATION OF A NOVEL PRIMATE T-CELL
; TITLE OF INVENTION: LYMPHOTROPIC VIRUS AND THE USE OF THIS VIRUS
; TITLE OF INVENTION: OR COMPONENTS THEREOF IN DIAGNOSTIC ASSAYS
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & PINNEGAN
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/04910
; FILING DATE: 21-APR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US08/231,526
; FILING DATE: 22-APR-1994

```
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: WILLIAM S. FEILER  
; REGISTRATION NUMBER: 26,728  
; REFERENCE/DOCKET NUMBER: 2026-4125PCT  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 758-4800  
; TELEX: (212) 751-6849  
; PCT-US95-04910-13
```

PCT-US95-04910-13

	Query Match	Best Local Similarity	Mismatches	Gaps	Indels	Length
QY	47 PRAPSLRRHYLIHNTLQQLOAALRLAPAPA-----LPPEPLFGEEDFSLSA	9.0%; Score 92.5; DB 5;	40;	Conservative 19;	Matches 48;	Indels 79;
DB	13 PKAP---RGLSTHHWLNFLQASYRLQGSDDFDQLRRFLKALKTPIWLNPIDYSLLA	21.5%; Pred. No. 0.99;				
QY	94 TI-----GSILRELDTSMDGTEPPQNVPVTLGLQNEVPQPDPVFLEALSRYLG					
DB	70 SLIPKGYPGRTSEIINVLRNQASPTPPPAPSLP-----EPANPPPL-----					
QY	145 SGLDDFFLDITSAVEKEPARAPP-----PHNLFCAFGSWEWNHE					
DB	112 -----QQPS-APPEHTPPPYIEPPATHCLPILHPHGAFSAHRPWQMKD					
QY	186 LDHIME					
DB	155 LQAIKQ					

```

RESULT 6
US-09-949-016-10257
; Sequence 10257, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10257
; LENGTH: 379
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-10257

```

	Query Match	8.9%;	Score 92;	DB 4;	Length 379;
	Best Local Similarity	21.6%;	Pred. No. 0.93;		
	Matches 44;	Conservative 23;	Mismatches 55;	Indels 82;	Gaps 10;
Qy	3	GGLKRKHSDLDEEEER-----	WESPAGLQSYQQALLRISLDKVQRSLGPR	48	
Dd	221	GGAAAPEDDDDDDEELLEVRLVETPGRELWRMPAG-----	RAARGQAERAQGGS	271	
Qy	49	APSLRRHVLHTNTLQQLQAALRLAPALPPEPLFLGEEDFSLSATIGSILRELDTSMDG	108		

Db 272 G-----EGAAAAAASPPTSEDEP--EEEEEEAAAEEG-----EETVASG 312

QY 109 TEPPQNPVTLGLQNEVPQPDPVFLEALSSRYLGDSDLDDFFLDIDTSAVEKEPARAPP 168

Db 313 EE-----SLGFLSLRLLP-----CPAGL-----DCSALDRDPDLQPP 343

QY 169 EPPHNL-----FCAP-----GSW 181

Db 344 SGTSHFEFPDYCTPEVTEMIAGDW 367

RESULT 7

US-09-377-285B-16

; Sequence 16, Application US/09377285B

; Patent No. 6720175

; GENERAL INFORMATION:

; APPLICANT: THE JOHNS HOPKINS UNIVERSITY SCHOOL OF MEDICINE

; APPLICANT: WORLEY, Paul

; APPLICANT: TU, Jian

; APPLICANT: XIAO, Bo

; APPLICANT: LEAHY, Daniel

; APPLICANT: BENEKEN, Jutta

; APPLICANT: LANAHAN, Anthony

; TITLE OF INVENTION: NUCLEIC ACID MOLECULE ENCODING HOMER 1b PROTEIN (AS AMENDED)

; FILE REFERENCE: JHU1580-4

; CURRENT APPLICATION NUMBER: US/09/377,285B

; CURRENT FILING DATE: 1999-08-18

; PRIOR APPLICATION NUMBER: US 60/138,426

; PRIOR FILING DATE: 1999-06-10

; PRIOR APPLICATION NUMBER: US 60/138,493

; PRIOR FILING DATE: 1999-06-10

; PRIOR APPLICATION NUMBER: US 60/138,494

; PRIOR FILING DATE: 1999-06-10

; PRIOR APPLICATION NUMBER: US 60/097,334

; PRIOR FILING DATE: 1998-08-18

; NUMBER OF SEQ ID NOS: 72

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 16

; LENGTH: 370

; TYPE: PRT

; ORGANISM: Rattus norvegicus

US-09-377-285B-16

Query Match 8.8%; Score 90.5; DB 4; Length 370;

Best Local Similarity 23.5%; Pred. No. 1.3;

Matches 43; Conservative 29; Mismatches 62; Indels 49; Gaps 8;

QY 33 LLRISLDKV-QRSLGPRAPSLRRHVLIHNTLQQLQAALRLAPALPP-----EPLFLG 85

Db 102 MVNMHLEKVARREIGTLA-----TVVRLPPSQKVIPPESLPPLTPYCRKPL--- 147

QY 86 EEDFSLSATIGSILRELDTSMDDT-----EPPQNPVTLGLQNEVPQ-PDPVFLEAL-- 137

Db 148 --NFACLDDVGHGVDLSTQLSRTGTLRSKSIKAPATPASATLGRPPRIPEVQLPAVPD 205

QY 138 -----SSRYLGDSDLDDFFLDIDTSAVEKEPARAPPEP-----PHNLFCAPG 179

Db 206 GKLSAASSVSSLASAGSAGSAGIPQSKGVAPATPPPPPIAPVTPPPPLPAEIFLLPP 265

QY 180 SWE 182

Db 266 PME 268

RESULT 8

US-08-887-518-2

; Sequence 2, Application US/08887518

; Patent No. 5843721

; GENERAL INFORMATION:

; APPLICANT: Rothe, Mike

; APPLICANT: Wu, Lin

; TITLE OF INVENTION: NIK Proteins, Nucleic Acids and Methods

; NUMBER OF SEQUENCES: 4

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP

; STREET: 268 BUSH STREET, SUITE 3200

; CITY: SAN FRANCISCO

; STATE: CALIFORNIA

; COUNTRY: USA

; ZIP: 94104

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/887,518

; FILING DATE:

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: OSMAN, RICHARD A

; REGISTRATION NUMBER: 36,627

; REFERENCE/DOCKET NUMBER: T97-008

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (415) 343-4341

; TELEFAX: (415) 343-4342

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 947 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

US-08-887-518-2

Query Match 8.8%; Score 90.5; DB 2; Length 947;

Best Local Similarity 25.8%; Pred. No. 4.3;

Matches 54; Conservative 17; Mismatches 87; Indels 51; Gaps 10;

QY 3 GGLKRRKHSLEEEERWEWSPAGLQSYQQALLRISLDKVQSLGPRAPSLRRHVLIHNTL 62

Db 659 GGLK---SPWRGEYKEPRHPPPNQANYHQ-----TLHAQPRELSRAPGPRPAEETTGRA 710

QY 63 QQLQAALRLAPALPPE-----PLFLGEEDFSLSATIGSILRELDTSMDDTTEPPQNPVT 117

Db 711 PKLQ-----PPLPPEPPEPNKSPPLTLSKEE-----SGMWEPLPLSSLEPAPARNPSS 758

QY 118 PLGLQNEVPQ-----PDPVFLEALSSRY-----LGDSGLDDFFLDIDTSAVEKEPA 164

Db 759 P-ERKATVPEQLQQLLELFLNLSLQPFSLQEEQILCLSIDSLSDDS---EKNPS 814

QY 165 RAPPEPPHNLFCAPGSW-----EWN 184

Db 815 KASQSSRDTLSSGVHWSWSSQAEARSSSWN 843

RESULT 9

US-09-023-321-2

; Sequence 2, Application US/09023321

; Patent No. 5844073

; GENERAL INFORMATION:

; APPLICANT: Rothe, Mike

; APPLICANT: Wu, Lin

; TITLE OF INVENTION: NIK Proteins, Nucleic Acids and Methods

; NUMBER OF SEQUENCES: 4

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP

; STREET: 268 BUSH STREET, SUITE 3200

; CITY: SAN FRANCISCO

; STATE: CALIFORNIA

; COUNTRY: USA

; ZIP: 94104

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/023,321
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/887,518
FILING DATE:

ATTORNEY/AGENT INFORMATION:
NAME: OSMAN, RICHARD A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: T97-008
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 343-4341
TELEFAX: (415) 343-4342
INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:
LENGTH: 947 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-023-321-2

Query Match 8.8%; Score 90.5; DB 2; Length 947;
Best Local Similarity 25.8%; Pred. No. 4.3;
Matches 54; Conservative 17; Mismatches 87; Indels 51; Gaps 10;

QY 3 GGLKRKHSDEEEERWESPAGLQSYQQALLRISLDKVQSLGPRAPSLRRHVLHNTL 62
Db 659 GGLK---SPWRGEYKEPRHPPNQANYHQ-----TLHAQPRELSPRAGPRAEETTGRA 710
QY 63 QQLQAALRLAPALPPE-----PLFLGEEDEFSLSATIGSILRELDTSMDGTEPPQNPVT 117
Db 711 PKLQ-----PPLPPEPPEPNKSPPLTLSKEE-----SGMWEPLPLSSLEPAPARNPSS 758
QY 118 PLGLQNEVPPQ-----PDPVFLEALSSRY-----LGDSGLDDFFLDIDTSAVEKEPA 164
Db 759 P-ERKATVPEQELQLEIEFLNLSLQPPFSLEEQEQLSCLSIDLSLSDDS---EKNPS 814
QY 165 RAPPEPPHNLFCAPGSW-----EWN 184
Db 815 KASQSSRDTLSSGVHSSWSSQAEARSSWN 843

RESULT 10
US-09-032-475-2
Sequence 2, Application US/09032475
Patent No. 5854003
GENERAL INFORMATION:
APPLICANT: Rothe, Mike
APPLICANT: Wu, Lin
TITLE OF INVENTION: NIK Proteins, Nucleic Acids and Methods
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
STREET: 268 BUSH STREET, SUITE 3200
CITY: SAN FRANCISCO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/032,475
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/887,518
FILING DATE:

ATTORNEY/AGENT INFORMATION:
NAME: OSMAN, RICHARD A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: T97-008
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 343-4341
TELEFAX: (415) 343-4342
INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:
LENGTH: 947 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-032-475-2

Query Match 8.8%; Score 90.5; DB 2; Length 947;
Best Local Similarity 25.8%; Pred. No. 4.3;
Matches 54; Conservative 17; Mismatches 87; Indels 51; Gaps 10;

QY 3 GGLKRKHSDEEEERWESPAGLQSYQQALLRISLDKVQSLGPRAPSLRRHVLHNTL 62
Db 659 GGLK---SPWRGEYKEPRHPPNQANYHQ-----TLHAQPRELSPRAGPRAEETTGRA 710
QY 63 QQLQAALRLAPALPPE-----PLFLGEEDEFSLSATIGSILRELDTSMDGTEPPQNPVT 117
Db 711 PKLQ-----PPLPPEPPEPNKSPPLTLSKEE-----SGMWEPLPLSSLEPAPARNPSS 758
QY 118 PLGLQNEVPPQ-----PDPVFLEALSSRY-----LGDSGLDDFFLDIDTSAVEKEPA 164
Db 759 P-ERKATVPEQELQLEIEFLNLSLQPPFSLEEQEQLSCLSIDLSLSDDS---EKNPS 814
QY 165 RAPPEPPHNLFCAPGSW-----EWN 184
Db 815 KASQSSRDTLSSGVHSSWSSQAEARSSWN 843

RESULT 11
US-09-257-703-1
Sequence 1, Application US/09257703
Patent No. 6265538
GENERAL INFORMATION:
APPLICANT: Greene, Warner C.
APPLICANT: Lin, Xin
APPLICANT: Gelezuinas, Romas
TITLE OF INVENTION: A NOVEL INHIBITOR OF THE INFLAMMATORY RESPONSE INDUCED
TITLE OF INVENTION: BY TNF-ALPHA AND IL-1
FILE REFERENCE: 30448.61USU1
CURRENT APPLICATION NUMBER: US/09/257,703
CURRENT FILING DATE: 1999-02-25
EARLIER APPLICATION NUMBER: 60/076,299
EARLIER FILING DATE: 1998-02-27
NUMBER OF SEQ ID NOS: 2
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1
LENGTH: 947
TYPE: PRT
ORGANISM: NF-KAPPA B INDUCING KINASE (NIK)
US-09-257-703-1

Query Match 8.8%; Score 90.5; DB 3; Length 947;
Best Local Similarity 25.8%; Pred. No. 4.3;
Matches 54; Conservative 17; Mismatches 87; Indels 51; Gaps 10;

QY 3 GGLKRKHSDEEEERWESPAGLQSYQQALLRISLDKVQSLGPRAPSLRRHVLHNTL 62
Db 659 GGLK---SPWRGEYKEPRHPPNQANYHQ-----TLHAQPRELSPRAGPRAEETTGRA 710
QY 63 QQLQAALRLAPALPPE-----PLFLGEEDEFSLSATIGSILRELDTSMDGTEPPQNPVT 117
Db 711 PKLQ-----PPLPPEPPEPNKSPPLTLSKEE-----SGMWEPLPLSSLEPAPARNPSS 758
QY 118 PLGLQNEVPPQ-----PDPVFLEALSSRY-----LGDSGLDDFFLDIDTSAVEKEPA 164

Db 759 P-ERKATVPEQELQLEIEFLFNLSQLSPFSLSEEQILSCLSLSDLSDDS---EKNPS 814

QY 165 RAPPEPPHNLFCAPGSW-----EWN 184

Db 815 KASQSSRDTLSSGVHSSWSSQAEARSSWN 843

RESULT 12

US-09-871-889A-1

; Sequence 1, Application US/09871889A

; Patent No. 6645728

; GENERAL INFORMATION:

; APPLICANT: Greene, Warner C.

; APPLICANT: Lin, Xin

; APPLICANT: Gelezuinas, Romas

; TITLE OF INVENTION: A NOVEL INHIBITOR OF THE INFLAMMATORY RESPONSE INDUCED BY TNF-ALFA

; FILE REFERENCE: 30448.61USD1

; CURRENT APPLICATION NUMBER: US/09/871,889A

; CURRENT FILING DATE: 2001-06-01

; PRIOR APPLICATION NUMBER: 09/257,703

; PRIOR FILING DATE: 1999-02-25

; PRIOR APPLICATION NUMBER: 60/076,299

; PRIOR FILING DATE: 1998-02-27

; NUMBER OF SEQ ID NOS: 10

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 1

; LENGTH: 947

; TYPE: PRT

; ORGANISM: NF-KAPPA B INDUCING KINASE (NIK)

US-09-871-889A-1

Query Match 8.8%; Score 90.5; DB 4; Length 947;

Best Local Similarity 25.8%; Pred. No. 4.3;

Matches 54; Conservative 17; Mismatches 87; Indels 51; Gaps 10;

QY 3 GGLKRKHSDLSEEEERWESPAGLQSYQQALLRISLDKVQSLGPRAPSLRRHVLIHNTL 62

Db 659 GGLK---SPWRGEYKEPRHPPPNQANYHQ-----TLHAQPRELSPRAPGPRPAEETGRA 710

QY 63 QQLQAALRLAPALPPE-----PLFLGEEEDFSLSATIGSILRELDTSMDCGTEPPQNPT 117

Db 711 PKLQ-----PPLPPEPEPNKSPPLTSLKEE-----SGMWEPLPLSLEPAPARNPSS 758

QY 118 PLGLQNEVPPQ-----PDPVFLEALSSRY-----LGDSGLDDFFLDIDTSAVEKEPA 164

Db 759 P-ERKATVPEQELQLEIEFLFNLSQLSPFSLSEEQILSCLSLSDLSDDS---EKNPS 814

QY 165 RAPPEPPHNLFCAPGSW-----EWN 184

Db 815 KASQSSRDTLSSGVHSSWSSQAEARSSWN 843

RESULT 13

US-09-902-540-11750

; Sequence 11750, Application US/09902540

; Patent No. 6833447

; GENERAL INFORMATION:

; APPLICANT: Goldman, Barry S.

; APPLICANT: Hinkle, Gregory J.

; APPLICANT: Slater, Steven C.

; APPLICANT: Wiegand, Roger C.

; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof

; FILE REFERENCE: 38-10(15849)B

; CURRENT APPLICATION NUMBER: US/09/902,540

; CURRENT FILING DATE: 2001-07-10

; PRIOR APPLICATION NUMBER: 60/217,883

; PRIOR FILING DATE: 2000-07-10

; NUMBER OF SEQ ID NOS: 16825

; SEQ ID NO 11750

; LENGTH: 1027

; TYPE: PRT

; ORGANISM: Myxococcus xanthus

US-09-902-540-11750

Query Match 8.8%; Score 90.5; DB 4; Length 1027;

Best Local Similarity 25.3%; Pred. No. 4.7;

Matches 42; Conservative 12; Mismatches 55; Indels 57; Gaps 7;

QY 36 ISLDKVQSLGPRAPSLRRHVLIHNTLQQLQAALRLAPALPPEPLFLGEEEDFSLSATI 95

Db 271 LDLSDAEAAAPPPRAP-----VAPKAPPV-----TSPAGS 301

QY 96 GSILRELDTSMDCGTEPPQNPTPLGLQNEVPPQDPV-----FLEALSSR----- 140

Db 302 GGIEFDLDAEDVSPPP--PVAPPPPRAPAAPPPAPVSAAGGVFEFDLSDDDAGTVPPPP 359

QY 141 -----YLGDSGLDDFFLDIDTSA-VEKEPARAPPEPPHNLFCAP 178

Db 360 PRAPSPVSAGGGSID---FDLDSLEDVEEAPPVPPPPPRAAVMPP 402

RESULT 14

US-08-778-717-9

; Sequence 9, Application US/08778717

; Patent No. 6602689

; GENERAL INFORMATION:

; APPLICANT: UENO, EIICHI

; APPLICANT: NOBUYUKI, FUJII

; APPLICANT: OKADA, MASAHISA

; TITLE OF INVENTION: FUSED DNA SEQUENCE, FUSED PROTEIN

; TITLE OF INVENTION: EXPRESSED FROM SAID FUSED DNA SEQUENCE AND METHOD FOR

; TITLE OF INVENTION: EXPRESSING SAID DNA SEQUENCE

; NUMBER OF SEQUENCES: 21

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,

; ADDRESSEE: P.C.

; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400

; CITY: ARLINGTON

; STATE: VA

; COUNTRY: USA

; ZIP: 22202

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/778,717

; FILING DATE: 12-DEC-1996

; CLASSIFICATION: 536

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: JP 352225/1995

; FILING DATE: 28-DEC-1995

; ATTORNEY/AGENT INFORMATION:

; NAME: OBLON, NORMAN F.

; REGISTRATION NUMBER: 24,618

; REFERENCE/DOCKET NUMBER: 2084-031-0

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 703-413-3000

; TELEFAX: 703-413-2220

; INFORMATION FOR SEQ ID NO: 9:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 215 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; ORIGINAL SOURCE:

; ORGANISM: RECOMBINANT

; PUBLICATION INFORMATION:

; AUTHORS: NOBUYUKI FUJII ET AL,

; TITLE: FUSED DNA SEQUENCE, FUSED PROTEIN EXPRESSED

; TITLE: FROM SAID FUSED DNA SEQUENCE AND METHOD OF

; TITLE: EXPRESSING SAID FUSED PROTEIN

```
; RELEVANT RESIDUES IN SEQ ID NO: 9: FROM 1 TO 215
US-08-778-717-9

Query Match      8.7%; Score 89.5; DB 4; Length 215;
Best Local Similarity 23.7%; Pred. No. 0.78;
Matches 41; Conservative 24; Mismatches 67; Indels 41; Gaps 6;

QY 2 EGGLKRKHSDEEE-----ERWESPAGLQSYQQALLRISLDKVQVSLG----- 46
Db 31 EGKAQPKVEIEDEELYNCAKEAMEACPVSAITIEAGGSSLVPRGSEFMGOIHGLSPTP 90

QY 47 -PRAPSLRRHVLIHNTLQQLQAALRLAPAPA-----LPPEPLFGEEDFSL 92
Db 91 IPKAP---RGLSTHWMFLQAAYRLQRPSPDFDQQLRRFLKALKTPINPIDYSLL 147

QY 93 ATI-----GSILRELDTSMDGTEPPQNVPVPLG---LQNEVPPQDPVFLEA 136
Db 148 ASLIPKGYPRVVEIINILVKQVSPSAPAAPVTPICPTTTTPPPPPPSPEA 200

RESULT 15
US-09-949-016-7139
; Sequence 7139, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7139
; LENGTH: 611
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-7139

Query Match      8.7%; Score 89.5; DB 4; Length 611;
Best Local Similarity 25.8%; Pred. No. 3;
Matches 48; Conservative 21; Mismatches 56; Indels 61; Gaps 9;

QY 12 LEEERWESPAGLQSYQQALLRISLDKVQVSLGPRAPSLRRHVLIHNTLQQLQAALRL 71
Db 435 LQKEQQR-----SKDLESRQR-----SLEQANRSLQLRIOELELQAQIH-----GLPV 477

QY 72 APAP-----ALPPEPLFGEEDFSLSATIGSILRELDTSMDGTEPPQNVPVPL 119
Db 478 PPTPGLLSLATTASDSLKPEQLDIEEGRPGAATFHV-----GGGPAQN----- 522

QY 120 GLQNEVPPQP--DPVFLEALSSRYLGDSDLDFDLIDTSAVEKE-----PA 164
Db 523 -APHQPPAPPSPDALLDHFPSDHLGLDLG-DPPHILGLEIDILMEEEGVVGGSLGGALSPL 580

QY 165 RAPPEP 170
Db 581 RAASDP 586
```

THIS PAGE BLANK (USPTO)

Db 181 ACCCTCCAACAGCTGCAGGCTGCACATTGCGCTGGCTCCCGCCCCCTGCCCTGCCCTCCCGGAG 240

QY 241 CCCCTCTTCTGGCGGAGGAGATTCTCCCTGTAGCCACCATTGGCTCTATCCTCAGG 300

Db 241 CCCCTCTTCTGGCGGAGGAGATTCTCCCTGTAGCCACCATTGGCTCTATCCTCAGG 300

QY 301 GAGCTGGACACCTCCATGGATGGGACTGAGCCCCCTCAGAAATCCAGTGACTCCCTTGGC 360

Db 301 GAGCTGGACACCTCCATGGATGGGACTGAGCCCCCTCAGAAATCCAGTGACTCCCTTGGC 360

QY 361 CTCCAGAAATGAAGTGCCACCCAGCCTGATCCAGTCTTCTTAGAAGCTCTGAGTCCCGG 420

Db 361 CTCCAGAAATGAAGTGCCACCCAGCCTGATCCAGTCTTCTTAGAAGCTCTGAGTCCCGG 420

QY 421 TACTTGGGGGACTCTGGCCTGGATGACTTCTTTCTGGACATTGACACATCTGCGGTAGAA 480

Db 421 TACTTGGGGGACTCTGGCCTGGATGACTTCTTTCTGGACATTGACACATCTGCGGTAGAA 480

QY 481 AAGGAGCCTGCACGGGCCCCACCAGAGCCTCCTCAACACCTCTTCTGTGCCCCAGGTTCT 540

Db 481 AAGGAGCCTGCACGGGCCCCACCAGAGCCTCCTCAACACCTCTTCTGTGCCCCAGGTTCT 540

QY 541 TGGGAGTGGAAATGAACCTGGATCACATCATGGAATCATTTCTGGGTCCTAA 591

Db 541 TGGGAGTGGAAATGAACCTGGATCACATCATGGAATCATTTCTGGGTCCTAA 591

RESULT 2

AF192529

LOCUS AF192529 Homo sapiens RPA-binding trans-activator (RBT1) mRNA linear PRI 26-OCT-2000

DEFINITION AF192529 subunit of replication protein A

ACCESSION AF192529

VERSION AF192529.1 GI:6180044

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 1542)

AUTHORS Cho,J.M., Song,D.J., Bergeron,J., Benlimame,N., Wold,M.S. and Alaoui-Jamali,M.A.

TITLE RBT1, a novel transcriptional co-activator, binds the second subunit of replication protein A

JOURNAL Nucleic Acids Res. 28 (18), 3478-3485 (2000)

MEDLINE 20440390

PUBMED 10982866

REFERENCE 2 (bases 1 to 1542)

AUTHORS Cho,J.M., Song,D.J. and Alaoui-Jamali,M.A.

TITLE Direct Submission

JOURNAL Submitted (06-OCT-1999) Experimental Medicine, Lady Davis Institute, 3999 Cote-Ste-Catherine, Montreal, QC H3T-1E2, Canada

FEATURES

source

1. .1542

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

gene

1. .1542

/gene="RBT1"

292. .882

/gene="RBT1"

/codon_start=1

/product="RPA-binding trans-activator"

/protein_id="AAF05761.1"

/db_xref="GI:6180045"

/translation="MEGGLKRKHSDLSEERWEWSWSPAGLQSYQQALLRISLDKQVRS LGPRAPSLRRHVLIHNTLQQLQAALRLAPALPPEPLFGEEDFSLSATIGSILREL DTSMGTEPPQNPVTPVLGLQNEVPPQDPVFLEALSSRYLGDGLDDFFLDIDTSAVE KEPARAPPEPHNLFCAPGSWENELDIMEIILGS"

ORIGIN

Query Match 100.0%; Score 591; DB 9; Length 1542;

Best Local Similarity 100.0%; Pred. No. 6.4e-120;

Matches 591; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGAGGGAGGCTTGAAGAGGAAACACTCTGATTTTGGAAGAGGAGGAGGAGGAGGAGGAGGAG 60

Db 292 ATGGAGGGAGGCTTGAAGAGGAAACACTCTGATTTTGGAAGAGGAGGAGGAGGAGGAGGAGGAG 351

QY 61 TGGAGTCCAGCAGCGCCTTCAGAGCTACCCAGCAAGCCCTGCTCCGCGCATCTCCCTAGACAAA 120

Db 352 TGGAGTCCAGCAGCGCCTTCAGAGCTACCCAGCAAGCCCTGCTCCGCGCATCTCCCTAGACAAA 411

QY 121 GTCCAGCGCAGCCTGGGCCCCCGAGCACCCAGCCTCCGAGGCATGTCTCATCCATAAC 180

Db 412 GTCCAGCGCAGCCTGGGCCCCCGAGCACCCAGCCTCCGAGGCATGTCTCATCCATAAC 471

QY 181 ACCCTCCAACAGCTGCAGGCTGCACCTTCGCTGGCTCCCGCCCCCTGCCCTGCCCTCCCGGAG 240

Db 472 ACCCTCCAACAGCTGCAGGCTGCACCTTCGCTGGCTCCCGCCCCCTGCCCTGCCCTCCCGGAG 531

QY 241 CCCCTCTTCTTCCGCGAGGAGGATTCTCCCTGTGCAGCCACCATTGGCTCTATCCTCAGG 300

Db 532 CCCCTCTTCTTCCGCGAGGAGGATTCTCCCTGTGCAGCCACCATTGGCTCTATCCTCAGG 591

QY 301 GAGCTGGACACCTTCCATGGATGGGACTGAGCCCCCTCAGAAATCCAGTGACTCCCTTGGC 360

Db 592 GAGCTGGACACCTTCCATGGATGGGACTGAGCCCCCTCAGAAATCCAGTGACTCCCTTGGC 651

QY 361 CTCCAGAAATGAAGTGCCACCCAGCCTGATCCAGTCTTCTTAGAAGCTCTGAGTCCCGG 420

Db 652 CTCCAGAAATGAAGTGCCACCCAGCCTGATCCAGTCTTCTTAGAAGCTCTGAGTCCCGG 711

QY 421 TACTTGGGGGACTCTGGCCTGGATGACTTCTTTCTGGACATTGACACATCTGCGGTAGAA 480

Db 712 TACTTGGGGGACTCTGGCCTGGATGACTTCTTTCTGGACATTGACACATCTGCGGTAGAA 771

QY 481 AAGGAGCCTGCACGGGCCCCACCAGAGCCTCCTCAACACCTCTTCTGTGCCCGAGTTCT 540

Db 772 AAGGAGCCTGCACGGGCCCCACCAGAGCCTCCTCAACACCTCTTCTGTGCCCGAGTTCT 831

QY 541 TGGGAGTGGAAATGAACCTGGATCACATCATGGAATCATTTCTGGGGTCCTAA 591

Db 832 TGGGAGTGGAAATGAACCTGGATCACATCATGGAATCATTTCTGGGGTCCTAA 882

RESULT 3

CQ783944

LOCUS CQ783944 Sequence 4084 from Patent EP1396543. linear PAT 17-MAR-2004

DEFINITION CQ783944

ACCESSION CQ783944

VERSION CQ783944.1 GI:45503845

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1

AUTHORS Ota,T., Nishikawa,T., Isogai,T., Hayashi,K., Ishii,S., Kawai,Y., Wakamatsu,A., Sugiyama,T., Nagai,K., Kojima,S., Otsuki,T. and Koga,H.

TITLE Primers for synthesizing full length cDNA clones and their use

JOURNAL Patent: EP 1396543-A 4084 10-MAR-2004;

Research Association for Biotechnology (JP)

FEATURES

source

1. .1325

/organism="Homo sapiens"

/mol_type="unassigned DNA"

/db_xref="taxon:9606"

93. .683

/note="unnamed protein product"

/codon_start=1

/protein_id="CAF86937.1"

/db_xref="GI:45503846"

/translation="MVGGLKRKHSDLSEERWEWSWSPAGLQSYQQALLRISLDKQVRS LGPRAPSLRRHVLIHNTLQQLQAALRLAPALPPEPLFGEEDFSLSATIGSILREL DTSMGTEPPQNPVTPVLGLQNEVPPQDPVFLEALSSRYLGDGLDDFFLDIDTSAVE

CDS

ORIGIN		KEPARAPPEPPHNLFCAPGSWEWNELDHIMEIILGS"	
Query Match		99.7%;	Score 589.4; DB 6; Length 1325;
Best Local Similarity		99.8%;	Pred. No. 1.5e-119;
Matches 590; Conservative		0;	Mismatches 1; Indels 0; Gaps 0;
QY	1	ATGGAGGGAGGCTTGAAGAGGAAACACTCTGATTTTGGAAAGAGGAGGAGGAGGTGGGAG	60
Db	93	ATGGTGGGAGGCTTGAAGAGGAAACACTCTGATTTTGGAAAGAGGAGGAGGAGGTGGGAG	152
QY	61	TGGAGTCCAGCAGGCTTTCAGAGCTACCAGCAAGCCCTGCTCCGCATCTCCCTAGACAAA	120
Db	153	TGGAGTCCAGCAGGCTTTCAGAGCTACCAGCAAGCCCTGCTCCGCATCTCCCTAGACAAA	212
QY	121	GTCCAGCGCAGCCTGGGCCCCCGAGCACCAGCCCTCCGAGGCATGTCTCTCATCCTAAAC	180
Db	213	GTCCAGCGCAGCCTGGGCCCCCGAGCACCAGCCCTCCGAGGCATGTCTCTCATCCTAAAC	272
QY	181	ACCCTCCAACAGCTGCAGGTGCACTTTCGCCCTGGCTCCGCCCTGCTGCCCTGCCCGGAG	240
Db	273	ACCCTCCAACAGCTGCAGGTGCACTTTCGCCCTGGCTCCGCCCTGCTGCCCTGCCCGGAG	332
QY	241	CCCCTCTTCTGGCGAGGAGATTCTCCCTGTCCAGCCACCATTTGGCTCTATCCTCAGG	300
Db	333	CCCCTCTTCTGGCGAGGAGATTCTCCCTGTCCAGCCACCATTTGGCTCTATCCTCAGG	392
QY	301	GAGCTGGACACCTCCATGGATGGGACTGAGCCCCCTCAGAATCCAGTCACTCCCTTGGC	360
Db	393	GAGCTGGACACCTCCATGGATGGGACTGAGCCCCCTCAGAATCCAGTCACTCCCTTGGC	452
QY	361	CTCCAGAATGAAGTGCACGGCCCCCAGCAGCCTCTCTTCTAGAACTCTGAGCTCCCGG	420
Db	453	CTCCAGAATGAAGTGCACGGCCCCCAGCAGCCTCTCTTCTAGAACTCTGAGCTCCCGG	512
QY	421	TACTTGGGGACTCTGGCCTGGATGACTTCTTTCTGGACATTGACACATCTCGGTAGAA	480
Db	513	TACTTGGGGACTCTGGCCTGGATGACTTCTTTCTGGACATTGACACATCTCGGTAGAA	572
QY	481	AAGGAGCCTGCACGGCCCCCAGCAGCCTCTCTCACAACCTCTTCTGTGCCCCAGGTTCT	540
Db	573	AAGGAGCCTGCACGGCCCCCAGCAGCCTCTCTCACAACCTCTTCTGTGCCCCAGGTTCT	632
QY	541	TGGGAGTGAATGAACCTGGATCACATCATGGAATCATTTCTGGGGTCTCTAA	591
Db	633	TGGGAGTGAATGAACCTGGATCACATCATGGAATCATTTCTGGGGTCTCTAA	683
RESULT 4		BD127913	
LOCUS		BD127913	
DEFINITION		Primer for synthesizing full-length cDNA and use thereof.	
ACCESSION		BD127913	
VERSION		BD127913.1	
KEYWORDS		JP 2002017375-A/3344.	
SOURCE		Homo sapiens (human)	
ORGANISM		Homo sapiens	
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
AUTHORS		1 (bases 1 to 1325)	
TITLE		Ota,T., Nishikawa,T., Isogai,T., Hayashi,K., Ishii,S., Kawai,Y., Wakamatsu,A., Sugiyama,T., Nagai,K., Kojima,S., Otsuki,T. and Koga,H.	
JOURNAL		Primer for synthesizing full-length cDNA and use thereof	
COMMENT		Patent: JP 2002017375-A 3344 22-JAN-2002;	
		HELIX RESEARCH INSTITUTE	
		OS Homo sapiens (human)	
		PN JP 2002017375-A/3344	
		PD 22-JAN-2002	
		PF 07-JUL-2000 JP 2000253172	
		PI TOSHIO OTA,TETSUO NISHIKAWA,TAKAO ISOGAI,KOJI HAYASHI,SHIZUKO	
		PI ISHII,	
		PI YURI KAWAI,AI WAKAMATSU,TOMOYASU SUGIYAMA,KEIICHI NAGAI, PI	
SHINICHI KOJIMA,		AK074604	
PI TETSUJI OTSUKI,HISASHI KOGA		1325 bp	
PC		mRNA	
C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/ PC		linear	
10,		PRI 03-SEP-2002	
PC C12P21/02,C12Q1/68//C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC		Homo sapiens cDNA FLJ90123 fis, clone HEMBA1007226, highly similar	
Primer for synthesizing full-length cDNA and use thereof FH Key		to Homo sapiens RPA-binding trans-activator (RBT1) mRNA.	
Location/Qualifiers		AK074604	
FT CDS	(93)..(680).	ACCESSION	
Location/Qualifiers		AK074604.1	
1..1325		VERSION	
/organism="Homo sapiens"		oligo capping; fis (full insert sequence).	
/mol_type="genomic DNA"		KEYWORDS	
/db_xref="taxon:9606"		SOURCE	
FEATURES		ORGANISM	
source		Homo sapiens	
ORIGIN		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
Query Match		99.7%;	
Best Local Similarity		99.8%;	
Matches 590; Conservative		0;	
Mismatches		1;	
Indels		0;	
Gaps		0;	
QY	1	ATGGAGGGAGGCTTGAAGAGGAAACACTCTGATTTTGGAAAGAGGAGGAGGAGGTGGGAG	60
Db	93	ATGGTGGGAGGCTTGAAGAGGAAACACTCTGATTTTGGAAAGAGGAGGAGGAGGTGGGAG	152
QY	61	TGGAGTCCAGCAGGCTTTCAGAGCTACCAGCAAGCCCTGCTCCGCATCTCCCTAGACAAA	120
Db	153	TGGAGTCCAGCAGGCTTTCAGAGCTACCAGCAAGCCCTGCTCCGCATCTCCCTAGACAAA	212
QY	121	GTCCAGCGCAGCCTGGGCCCCCGAGCACCAGCCCTCCGAGGCATGTCTCTCATCCTAAAC	180
Db	213	GTCCAGCGCAGCCTGGGCCCCCGAGCACCAGCCCTCCGAGGCATGTCTCTCATCCTAAAC	272
QY	181	ACCCTCCAACAGCTGCAGGTGCACTTTCGCCCTGGCTCCGCCCTGCTGCCCTGCCCGGAG	240
Db	273	ACCCTCCAACAGCTGCAGGTGCACTTTCGCCCTGGCTCCGCCCTGCTGCCCTGCCCGGAG	332
QY	241	CCCCTCTTCTGGCGAGGAGATTCTCCCTGTCCAGCCACCATTTGGCTCTATCCTCAGG	300
Db	333	CCCCTCTTCTGGCGAGGAGATTCTCCCTGTCCAGCCACCATTTGGCTCTATCCTCAGG	392
QY	301	GAGCTGGACACCTCCATGGATGGGACTGAGCCCCCTCAGAATCCAGTCACTCCCTTGGC	360
Db	393	GAGCTGGACACCTCCATGGATGGGACTGAGCCCCCTCAGAATCCAGTCACTCCCTTGGC	452
QY	361	CTCCAGAATGAAGTGCACGGCCCCCAGCAGCCTCTCTTCTAGAACTCTGAGCTCCCGG	420
Db	453	CTCCAGAATGAAGTGCACGGCCCCCAGCAGCCTCTCTTCTAGAACTCTGAGCTCCCGG	512
QY	421	TACTTGGGGACTCTGGCCTGGATGACTTCTTTCTGGACATTGACACATCTCGGTAGAA	480
Db	513	TACTTGGGGACTCTGGCCTGGATGACTTCTTTCTGGACATTGACACATCTCGGTAGAA	572
QY	481	AAGGAGCCTGCACGGCCCCCAGCAGCCTCTCTCACAACCTCTTCTGTGCCCCAGGTTCT	540
Db	573	AAGGAGCCTGCACGGCCCCCAGCAGCCTCTCTCACAACCTCTTCTGTGCCCCAGGTTCT	632
QY	541	TGGGAGTGAATGAACCTGGATCACATCATGGAATCATTTCTGGGGTCTCTAA	591
Db	633	TGGGAGTGAATGAACCTGGATCACATCATGGAATCATTTCTGGGGTCTCTAA	683
RESULT 5		AK074604	
LOCUS		AK074604	
DEFINITION		Homo sapiens cDNA FLJ90123 fis, clone HEMBA1007226, highly similar	
ACCESSION		AK074604	
VERSION		AK074604.1	
KEYWORDS		oligo capping; fis (full insert sequence).	
SOURCE		Homo sapiens (human)	
ORGANISM		Homo sapiens	
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
AUTHORS		1 (bases 1 to 1325)	
TITLE		Ota,T., Nishikawa,T., Isogai,T., Hayashi,K., Ishii,S., Kawai,Y., Wakamatsu,A., Sugiyama,T., Nagai,K., Kojima,S., Otsuki,T. and Koga,H.	
JOURNAL		Primer for synthesizing full-length cDNA and use thereof	
COMMENT		Patent: JP 2002017375-A 3344 22-JAN-2002;	
		HELIX RESEARCH INSTITUTE	
		OS Homo sapiens (human)	
		PN JP 2002017375-A/3344	
		PD 22-JAN-2002	
		PF 07-JUL-2000 JP 2000253172	
		PI TOSHIO OTA,TETSUO NISHIKAWA,TAKAO ISOGAI,KOJI HAYASHI,SHIZUKO	
		PI ISHII,	
		PI YURI KAWAI,AI WAKAMATSU,TOMOYASU SUGIYAMA,KEIICHI NAGAI, PI	

REFERENCE 1 Isogai,T., Ota,T., Nishikawa,T., Hayashi,K., Otsuki,T., Sugiyama,T., Suzuki,Y., Nagai,K., Sugano,S., Ishii,S., Kawai-Hio,Y., Saito,K., Yamamoto,J., Wakamatsu,A., Nakamura,Y., Kojima,S., Nagahari,K., Masuho,Y., Ono,T., Okano,K., Yoshikawa,Y., Aotsuka,S., Sasaki,N., Hattori,A., Okumura,K., Iwayanagi,T. and Ninomiya,K.

TITLE NEDO human cDNA sequencing project

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 1325)

AUTHORS Isogai,T. and Otsuki,T.

TITLE Direct Submission

JOURNAL Submitted (25-MAR-2002) Takao Isogai, Helix Research Institute, Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)

COMMENT NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology; cDNA library construction: Institute of Medical Science, University of Tokyo, Laboratory of Genome Structure, Human Genome Center; cDNA 5'- & 3'-end one pass sequencing and clone selection: Helix Research Institute (supported by Japan Key Technology Center etc.).

FEATURES source Location/Qualifiers 1. .1325 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="HEMBA1007226" /tissue_type="whole embryo, mainly head" /clone_lib="HEMBA1" /dev_stages="embryo, 10 weeks" /note="cloning vector: pME18SFL3"

ORIGIN Query Match 99.7%; Score 589.4; DB 9; Length 1325; Best Local Similarity 99.8%; Pred. No. 1.5e-119; Matches 590; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGAGGGAGGCTTGAAGAGGAAACACTCTGATTTTGAAGAGGAGGAGGAGGTGGGAG 60

Db 93 ATGGTGGGAGGCTTGAAGAGGAAACACTCTGATTTTGAAGAGGAGGAGGAGGTGGGAG 152

QY 61 TGGAGTCCAGCAGGCCCTTCAGAGCTACCAGCAAGCCCTGCTCCGCATCTCCCTAGACAAA 120

Db 153 TGGAGTCCAGCAGGCCCTTCAGAGCTACCAGCAAGCCCTGCTCCGCATCTCCCTAGACAAA 212

QY 121 GTCCAGCGCAGCCTGGGCCCCCGAGCACCACCCAGCCTCCGCAGGCATGTCCTCATCCATAAC 180

Db 213 GTCCAGCGCAGCCTGGGCCCCCGAGCACCACCCAGCCTCCGCAGGCATGTCCTCATCCATAAC 272

QY 181 ACCCTCCAACAGCTGCAGGTGCACATTTCGCCTGGCTCCGCCCCCTGCCCTGCCCCCGAG 240

Db 273 ACCCTCCAACAGCTGCAGGTGCACATTTCGCCTGGCTCCGCCCCCTGCCCTGCCCCCGAG 332

QY 241 CCCCTCTTCCTGGCGAGGAGGATTTCTCCCTGTCAGCCACCATTTGGCTCTATCCTCAGG 300

Db 333 CCCCTCTTCCTGGCGAGGAGGATTTCTCCCTGTCAGCCACCATTTGGCTCTATCCTCAGG 392

QY 301 GAGCTGGACACCTCCATGGATGGGACTGAGCCCCCTCAGAATCCAGTGACTCCCTTGGC 360

Db 393 GAGCTGGACACCTCCATGGATGGGACTGAGCCCCCTCAGAATCCAGTGACTCCCTTGGC 452

QY 361 CTCCAGAATGAAGTGCACCCAGCCCTGATCCAGTCTTCTTAGAAGCTCTGAGTCCCGG 420

Db 453 CTCCAGAATGAAGTGCACCCAGCCCTGATCCAGTCTTCTTAGAAGCTCTGAGTCCCGG 512

QY 421 TACTTGGGGGACTCTGGCCTGGATGACTTCTTTCTGGACATTGACACATCTGCGGTAGAA 480

Db 513 TACTTGGGGGACTCTGGCCTGGATGACTTCTTTCTGGACATTGACACATCTGCGGTAGAA 572

QY 481 AAGGAGCCTGCACGGGCCCCACGAGCCCTCCTCAACACCTCTTCTGTGCCCCCAGGTTCT 540

Db 573 AAGGAGCCTGCACGGGCCCCACGAGCCCTCCTCAACACCTCTTCTGTGCCCCCAGGTTCT 632

QY 541 TGGGAGTGAATGAACTGGATCACATCATGGAATCATTTCTGGGGTCCTAA 591

Db 633 TGGGAGTGAATGAACTGGATCACATCATGGAATCATTTCTGGGGTCCTAA 683

RESULT 6

BC014061 1346 bp mRNA linear PRI 29-JUN-2004

LOCUS BC014061

DEFINITION Homo sapiens SERTA domain containing 3, transcript variant 2, mRNA (cDNA clone MGC:19994 IMAGE:3926937), complete cds.

ACCESSION BC014061

VERSION BC014061.2 GI:33873834

KEYWORDS MGC:

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 1346)

REFERENCE Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D., Altshul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Kettman,M., Madan,A., Young,A.C., Shevchenko,Y., Sanchez,A., Whiting,M., Madan,A., Touchman,J.W., Green,E.D., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butlerfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E., Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.

TITLE Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

PUBMED 12477932

REFERENCE 2 (bases 1 to 1346)

AUTHORS Strausberg,R.

TITLE Direct Submission

JOURNAL Submitted (10-SEP-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>

COMMENT On Aug 19, 2003 this sequence version replaced gi:15559388. Contact: MGC help desk Email: cgapbs-remail.nih.gov Tissue Procurement: DCTD/DTP CDNA Library Preparation: Rubin Laboratory CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland; Web site: <http://www.nisc.nih.gov/> Contact: nisc_mgc@nhgri.nih.gov Akhter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B., Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S., Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P., Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Laric,P., Legaspi,R., Maduro,Q.L., Masiello,C., Maskeri,B., Mastrian,S.D., McCloskey,J.C., McDowell,J., Pearson,R., Stantripop,S., Thomas,P.J., Touchman,J.W., Tsurgeon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L., Young,A., Zhang,L.-H. and Green,E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov> Series: IRAL Plate: 28 Row: f Column: 10 This clone was selected for full length sequencing because it

FEATURES	passed the following selection criteria: matched mRNA gi: 15718683.			RESULT 7	
	source	Location/Qualifiers		BC050643	
gene		1. .1346		LOCUS	
		/organism="Homo sapiens"		DEFINITION	Homo sapiens SERTA domain containing 3, transcript variant 2, mRNA (CDNA clone MGC:60134 IMAGE:6503515), complete cds.
		/mol_type="mRNA"		ACCESSION	BC050643
		/db_xref="taxon:9606"		VERSION	BC050643.2
		/clone="MGC:19994 IMAGE:3926937"		KEYWORDS	GI:34193422
		/tissue type="Ovary, adenocarcinoma"		SOURCE	MGC
		/clone_lib="NIH MGC_9"		ORGANISM	Homo sapiens (human)
		/lab_host="DH10B-R"			Homo sapiens
		/note="Vector: pOTB7"			Homo sapiens
		1. .1346		REFERENCE	1 (bases 1 to 1420)
CDS		/gene="SERTAD3"		AUTHORS	Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D., Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E., Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
		/db_xref="GI:15559389"		TITLE	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
		/db_xref="LocusID:29946"		JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
		/translation="MVGGLKRKHSDEEEERWESPAGLQSYQQALLRISLDKVQRS LGPRAPSLRRHVLHNTLQQLQALRLAPALPPEPLFLGEEDFSLSATIGSILREL DTSMDGTEPPQNPVTPGLQNEVPPQDPVFLEALSSRYLGDSDLDDFFLDIDTSAVE KEPARAPPEPPHNLFCAPGSWEWNELDHIMEILGS"		REFERENCE	12477932
		/product="RPA-binding trans-activator"		AUTHORS	2 (bases 1 to 1420)
		/protein_id="AAH14061.1"		TITLE	Strausberg,R.
		/db_xref="GI:15559389"		JOURNAL	Direct Submission
		/db_xref="LocusID:29946"			Submitted (08-APR-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
		/codon_start=1		REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov
		/note="SERTAD3"		COMMENT	On Aug 25, 2003 this sequence version replaced gi:29791864.
ORIGIN		/note="synonym: RBT1"			Contact: MGC help desk
		/db_xref="LocusID:29946"			Email: cgapbs-r@mail.nih.gov
		92_ .682			Tissue Procurement: ATCC
		/gene="SERTAD3"			CDNA Library Preparation: Life Technologies, Inc.
		/codon_start=1			CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
		/product="RPA-binding trans-activator"			DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305
		/protein_id="AAH14061.1"			Web site: http://www-shgc.stanford.edu
		/db_xref="GI:15559389"			Contact: (Dickson, Mark) mcd@paxil.stanford.edu
		/db_xref="LocusID:29946"			Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.
		/translation="MVGGLKRKHSDEEEERWESPAGLQSYQQALLRISLDKVQRS LGPRAPSLRRHVLHNTLQQLQALRLAPALPPEPLFLGEEDFSLSATIGSILREL DTSMDGTEPPQNPVTPGLQNEVPPQDPVFLEALSSRYLGDSDLDDFFLDIDTSAVE KEPARAPPEPPHNLFCAPGSWEWNELDHIMEILGS"			
Query Match	99.7%;	Score 589.4;	DB 9;	Length 1346;	
Best Local Similarity	99.8%;	Pred. No. 1.5e-119;			
Matches	590;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;
QY	1	ATGGAGGGAGGCTTGAAGAGGAAACACTCTGATTTCGAAGAGGAGGAGGAGGTGGGAG	60		
Db	92	ATGGTGGAGGCTTGAAGAGGAAACACTCTGATTTCGAAGAGGAGGAGGAGGTGGGAG	151		
QY	61	TGGAGTCCAGCAGGCCCTTCAGAGCTACCAGCAAGCCCTGCTCCGCATCTCCCTAGACAAA	120		
Db	152	TGGAGTCCAGCAGGCCCTTCAGAGCTACCAGCAAGCCCTGCTCCGCATCTCCCTAGACAAA	211		
QY	121	GTCCAGCGCAGCCTGGGCCCCCGAGCACCCAGCCCTCCGAGGCATGTCCTCATCCATAAC	180		
Db	212	GTCCAGCGCAGCCTGGGCCCCCGAGCACCCAGCCCTCCGAGGCATGTCCTCATCCATAAC	271		
QY	181	ACCCTCCAACAGCTGCAGGTGCACCTTCGCCTGGCTCCCGCCCTGCCCTGCCCTCCCGGAG	240		
Db	272	ACCCTCCAACAGCTGCAGGTGCACCTTCGCCTGGCTCCCGCCCTGCCCTGCCCTCCCGGAG	331		
QY	241	CCCCTCTCTCTGGCGGAGGAGATTCTCCCTGTCAAGCACCATTGGCTCTATCTCAGG	300		
Db	332	CCCCTCTCTCTGGCGGAGGAGATTCTCCCTGTCAAGCACCATTGGCTCTATCTCAGG	391		
QY	301	GAGCTGGACACCTCCATGGATGGGACTGAGCCCCCTCAGAATCCAGTGACTCCCTTGGC	360		
Db	392	GAGCTGGACACCTCCATGGATGGGACTGAGCCCCCTCAGAATCCAGTGACTCCCTTGGC	451		
QY	361	CTCCAGAAATGAAGTGCCACCCAGCCCTGATCCAGTCTTCTTAGAAGCTCTGAGCTCCCGG	420		
Db	452	CTCCAGAAATGAAGTGCCACCCAGCCCTGATCCAGTCTTCTTAGAAGCTCTGAGCTCCCGG	511		
QY	421	TACTTGGGGACTCTGGCTGGATGACTTCTTTCTTGACATTGACACATCTGGGGTAGAA	480		
Db	512	TACTTGGGGACTCTGGCTGGATGACTTCTTTCTTGACATTGACACATCTGGGGTAGAA	571		
QY	481	AAGGAGCCTGCACGGGCCCCAGCAGCCCTCCTCAACAACCTTCTTGTCGCCCCAGGTTCT	540		
Db	572	AAGGAGCCTGCACGGGCCCCAGCAGCCCTCCTCAACAACCTTCTTGTCGCCCCAGGTTCT	631		
QY	541	TGGGAGTGAATGAATGAATGCATCATCATGGAATCATTTCTGGGTCCTAA	591		
Db	632	TGGGAGTGAATGAATGAATGCATCATCATGGAATCATTTCTGGGTCCTAA	682		

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 110 Row: o Column: 24
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 15718683.

FEATURES
source
1. .1420
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:60134 IMAGE:6503515"
/tissue type="Uterus, leiomyosarcoma"
/clone_lib="NIH MGC_71"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
1. .1420
/gene="SERTAD3"

gene

/note="synonym: RBT1"
 /db_xref="LocusID:29946"
 168. .758
 /gene="SERTAD3"
 /codon_start=1
 /product="RPA-binding trans-activator"
 /protein_id="AAH50643.1"
 /db_xref="GI:29791865"
 /db_xref="LocusID:29946"
 /translation="MVGGLKRKHSDLSEEEERWESPAGLSYQQALLRISLDKVKQRS
 LGPRAPSLRRHVLHNTLQQLAALRLAPALPPEPLFGEEDFSLSATIGSILREL
 DTSMDGTEPPQNVPVPLGLQNEVPPQDPVFLEALSRYLGDGLDDDFLDIDTSAVE
 KEPARAPPEPPHNLFCAPGSWEWNELDHIMEILGS"

ORIGIN

	Query Match	99.7%;	Score 589.4;	DB 9;	Length 1420;
	Best Local Similarity	99.8%;	Pred. No. 1.5e-119;		
	Matches 590;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;
QY	1	ATGGAGGGAGGCTTGAAGAGGAAACACTCTGATTTGGAAAGAGGAGGAGGAGAGGTGGGAG	60		
Db	168	ATGGTGGAGGCTTGAAGAGGAAACACTCTGATTTGGAAAGAGGAGGAGGAGGAGGTGGGAG	227		
QY	61	TGGAGTCCAGCAGGCCCTTCAGAGCTACCAAGCAAGCCCTGCTCCGCATCTCCCTAGACAAA	120		
Db	228	TGGAGTCCAGCAGGCCCTTCAGAGCTACCAAGCAAGCCCTGCTCCGCATCTCCCTAGACAAA	287		
QY	121	GTCCAGCGCAGCCTGGGCCCCCGAGCACCCAGCCTCCGAGGCATGTCTCATCCATAAC	180		
Db	288	GTCCAGCGCAGCCTGGGCCCCCGAGCACCCAGCCTCCGAGGCATGTCTCATCCATAAC	347		
QY	181	ACCCCTCCAAACAGCTGCAGGCTGCACCTTCGCTGGCTCCCGCCCCCTGCCCTGCCCCCGGAG	240		
Db	348	ACCCCTCCAAACAGCTGCAGGCTGCACCTTCGCTGGCTCCCGCCCCCTGCCCTGCCCCCGGAG	407		
QY	241	CCCCCTCTTCTGGGCGAGGAGGATTTCTCCCTGTGAGCCACCATTTGGCTCTATCTCTCAGG	300		
Db	408	CCCCCTCTTCTGGGCGAGGAGGATTTCTCCCTGTGAGCCACCATTTGGCTCTATCTCTCAGG	467		
QY	301	GAGCTGGACACCTCCATGGATGGGACTGAGCCCCCTCAGAAATCCAGTGACTCCCCCTTGGC	360		
Db	468	GAGCTGGACACCTCCATGGATGGGACTGAGCCCCCTCAGAAATCCAGTGACTCCCCCTTGGC	527		
QY	361	CTCCAGAAATGAAGTGCCACCCAGCCTGATCCAGTCTTCTTAGAAGCTCTGAGCTCCCCG	420		
Db	528	CTCCAGAAATGAAGTGCCACCCAGCCTGATCCAGTCTTCTTAGAAGCTCTGAGCTCCCCG	587		
QY	421	TACTTGGGSACTCTGGCCTGGATGACTTCTTTCTGGACATTGACACATCTGCGGTAGAA	480		
Db	588	TACTTGGGSACTCTGGCCTGGATGACTTCTTTCTGGACATTGACACATCTGCGGTAGAA	647		
QY	481	AAGGAGCCTGCACGGGCCCCCACCAGAGCCCTCCTCACAACCTCTTCTGTGCCCAGGTTCT	540		
Db	648	AAGGAGCCTGCACGGGCCCCCACCAGAGCCCTCCTCACAACCTCTTCTGTGCCCAGGTTCT	707		
QY	541	TGGGAGTGGAAATGAAGTGGATCACATCATGGAATCATTTCTGGGGTCTTAA	591		
Db	708	TGGGAGTGGAAATGAAGTGGATCACATCATGGAATCATTTCTGGGGTCTTAA	758		

RESULT 8
 LOCUS
 CQ731553
 DEFINITION
 Sequence 17487 from Patent WO02068579.
 ACCESSION
 CQ731553
 VERSION
 CQ731553.1 GI:42308440
 KEYWORDS
 SOURCE
 Homo sapiens (human)
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1
 Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.

AUTHORS DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL Submitted (15-SEP-1999) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE 3 (bases 1 to 160643)
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
JOURNAL Direct Submission
REFERENCE 4 (bases 1 to 160643)
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Submitted (03-OCT-2001) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE 5 (bases 1 to 160643)
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Submitted (13-JUL-2002) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT On Jul 13, 2002 this sequence version replaced gi:15887272.
Draft Sequence Produced by DOE Joint Genome Institute
www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center
www-shgc.stanford.edu
Quality: Phrap Quality >=40 100% of Sequence;
Estimated Total Number of Errors is 0.1.
FEATURES
source
1. .160643
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="19"
/clone="CTC-492K19"
ORIGIN
Query Match 99.7%; Score 589.4; DB 9; Length 160643;
Best Local Similarity 99.8%; Pred. No. 7.9e-120;
Matches 590; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGGAGGGAGGCTTGAAGAGGAAACACTCTGATTGGAAGAGGAGGAGGAGGAGGTGGGAG 60
Db 137453 ATGGTGGGAGGCTTGAAGAGGAAACACTCTGATTGGAAGAGGAGGAGGAGGAGGTGGGAG 137394
QY 61 TGGAGTCCAGCAGGCTTCAGAGCTACCCAGCAAGCCCTGCTCCGCATCTCCCTAGACAAA 120
Db 137393 TGGAGTCCAGCAGGCTTCAGAGCTACCCAGCAAGCCCTGCTCCGCATCTCCCTAGACAAA 137334
QY 121 GTCCAGCGCAGCCTGGSCCCCGAGCACCAGCCCTCCGCAGGCATGTCCTCATCCATAAC 180
Db 137333 GTCCAGCGCAGCCTGGSCCCCGAGCACCAGCCCTCCGCAGGCATGTCCTCATCCATAAC 137274
QY 181 ACCCTCCAACAGCTGAGGCTGCACCTTGGCTGCCGCTCCGCCCTGCCCTGCCCTGCCCTGCCGAG 240
Db 137273 ACCCTCCAACAGCTGAGGCTGCACCTTGGCTGCCGCTCCGCCCTGCCCTGCCCTGCCCTGCCGAG 137214
QY 241 CCCCTCTTCTGGCGAGGAGGATTCTCCCTGTGTCAGCCACCATTTGGCTCTATCCTCAGG 300
Db 137213 CCCCTCTTCTGGCGAGGAGGATTCTCCCTGTGTCAGCCACCATTTGGCTCTATCCTCAGG 137154
QY 301 GAGCTGGACACCTCCATGGATGGGACTGAGCCCTCAGAATCCAGTGACTCCCTTGGC 360
Db 137153 GAGCTGGACACCTCCATGGATGGGACTGAGCCCTCAGAATCCAGTGACTCCCTTGGC 137094
QY 361 CTCCAGAAATGAAGTGCACCCCGAGCCTGATCCAGTCTTTAGAAAGCTCTGAGCTCCCGG 420
Db 137093 CTCCAGAAATGAAGTGCACCCCGAGCCTGATCCAGTCTTTAGAAAGCTCTGAGCTCCCGG 137034
QY 421 TACTTGGGGGACTCTGGCCTGGATGACTTCTTTCTGGACATTGACACATCTGCGGTAGAA 480
Db 137033 TACTTGGGGGACTCTGGCCTGGATGACTTCTTTCTGGACATTGACACATCTGCGGTAGAA 136974
QY 481 AAGGAGCCTGCACGGGCCCCAGCAGGCTCTCTCAAAACCTCTTCTGTGCCCGAGGTTCT 540

Db 136973 AAGGAGCCTGCACGGGCCCCACCAGAGCCTCTCTCACAACCTCTTCTGTGCCCGAGGTTCT 136914
QY 541 TGGGAGTGAATGAAGTGGATCAGATCATCATGGAATCATTTCTGGGGTCTCTAA 591
Db 136913 TGGGAGTGAATGAAGTGGATCAGATCATCATGGAATCATTTCTGGGGTCTCTAA 136863
RESULT 10
CQ779981
LOCUS CQ779981 825 bp DNA linear PAT 17-MAR-2004
DEFINITION Sequence 121 from Patent EP1396543.
ACCESSION CQ779981
VERSION CQ779981.1 GI:45536053
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Ota, T., Nishikawa, T., Isogai, T., Hayashi, K., Ishii, S., Kawai, Y., Wakamatsu, A., Sugiyama, T., Nagai, K., Kojima, S., Otsuki, T. and Koga, H.
TITLE Primers for synthesizing full length cDNA clones and their use
JOURNAL Patent: EP 1396543-A 121 10-MAR-2004;
Research Association for Biotechnology (JP)
FEATURES
source
1. .825
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 99.6%; Score 588.4; DB 6; Length 825;
Best Local Similarity 99.7%; Pred. No. 2.6e-119;
Matches 589; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 ATGGAGGGAGGCTTGAAGAGGAAACACTCTGATTGGAAGAGGAGGAGGAGGAGGTGGGAG 60
Db 93 ATGGTGGGAGGCTTGAAGAGGAAACACTCTGATTGGAAGAGGAGGAGGAGGAGGTGGGAG 152
QY 61 TGGAGTCCAGCAGGCTTCAGAGCTACCCAGCAAGCCCTGCTCCGCATCTCCCTAGACAAA 120
Db 153 TGGAGTCCAGCAGGCTTCAGAGCTACCCAGCAAGCCCTGCTCCGCATCTCCCTAGACAAA 212
QY 121 GTCCAGCGCAGCCTGGSCCCCGAGCACCAGCCCTCCGCAGGCATGTCCTCATCCATAAC 180
Db 213 GTCCAGCGCAGCCTGGSCCCCGAGCACCAGCCCTCCGCAGGCATGTCCTCATCCATAAC 272
QY 181 ACCCTCCAACAGCTGAGGCTGCACCTTGGCTGCCGCTCCGCCCTGCCCTGCCCTGCCGAG 240
Db 273 ACCCTCCAACAGCTGAGGCTGCACCTTGGCTGCCGCTCCGCCCTGCCCTGCCCTGCCGAG 332
QY 241 CCCTCTTCTTCTGGCGAGGAGGATTCTCCCTGTGTCAGCCACCATTTGGCTCTATCCTCAGG 300
Db 333 CCCTCTTCTTCTGGCGAGGAGGATTCTCCCTGTGTCAGCCACCATTTGGCTCTATCCTCAGG 392
QY 301 GAGCTGGACACCTCCATGGATGGGACTGAGCCCTCAGAATCCAGTGACTCCCTTGGC 360
Db 393 GAGCTGGACACCTCCATGGATGGGACTGAGCCCTCAGAATCCAGTGACTCCCTTGGC 452
QY 361 CTCCAGAAATGAAGTGCACCCCGAGCCTGATCCAGTCTTTAGAAAGCTCTGAGCTCCCGG 420
Db 453 CTCCAGAAATGAAGTGCACCCCGAGCCTGATCCAGTCTTTAGAAAGCTCTGAGCTCCCGG 512
QY 421 TACTTGGGGGACTCTGGCCTGGATGACTTCTTTCTGGACATTGACACATCTGCGGTAGAA 480
Db 513 TACTTGGGGGACTCTGGCCTGGATGACTTCTTTCTGGACATTGACACATCTGCGGTAGAA 572
QY 481 AAGGAGCCTGCACGGGCCCCAGCAGGCTCTCTCAAAACCTCTTCTGTGCCCGAGGTTCT 540
Db 573 AAGGAGCCTGCACGGGCCCCAGCAGGCTCTCTCAAAACCTCTTCTGTGCCCGAGGTTCT 632
QY 541 TGGGAGTGAATGAAGTGGATCAGATCATCATGGAATCATTTCTGGGGTCTCTAA 591

Db 633 TGGGAGTGGAAATGAACCTGGATCACAATCATGGAATCATTTCTGGGTCCTAA 683
|||||
RESULT 11
CQ781937
LOCUS CQ781937 825 bp DNA linear PAT 17-MAR-2004
DEFINITION Sequence 2077 from Patent EP1396543.
ACCESSION CQ781937
VERSION CQ781937.1 GI:45537993
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Ota,T., Nishikawa,T., Isogai,T., Hayashi,K., Ishii,S., Kawai,Y.,
Wakamatsu,A., Sugiyama,T., Nagai,K., Kojima,S., Otsuki,T. and
Koga,H.
TITLE Primers for synthesizing full length cDNA clones and their use
JOURNAL Patent: EP 1396543-A 2077 10-MAR-2004;
Research Association for Biotechnology (JP)
FEATURES
source 1..825
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 99.6%; Score 588.4; DB 6; Length 825;
Best Local Similarity 99.7%; Pred. No. 2.6e-119;
Matches 589; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 ATGGAGGGAGGCTTGAAGAGGAACACTCTGATTTGGAAGAGGAGGAGGAGGTGGGAG 60
|||
Db 93 ATGGTGGAGGCTTGAAGAGGAACACTCTGATTTGGAAGAGGAGGAGGAGGTGGGAG 152
|||
QY 61 TGGAGTCCAGCAGGCCCTTCAGAGCTACCAGCAAGCCCTGCTCCGCATCTCCCTAGACAAA 120
|||
Db 153 TGGAGTCCAGCAGGCCCTTCAGAGCTACCAGCAAGCCCTGCTCCGCATCTCCCTAGACAAA 212
|||
QY 121 GTCCAGCGCAGCCTGGGCCCCCGAGCACCCAGCCCTCCGCAGGCATGTCCTCATCCATAAC 180
|||
Db 213 GTCCAGCGCAGCCTGGGCCCCCGAGCACCCAGCCCTCCGCAGGCATGTCCTCATCCATAAC 272
|||
QY 181 ACCCTCCAACAGCTGCAGGCTGCACCTTCGCCTGGCTCCGCCCTGCCCTGCCCTGCCCTGCCG 240
|||
Db 273 ACCCTCCAACAGCTGCAGGCTGCACCTTCGCCTGGCTCCGCCCTGCCCTGCCCTGCCCTGCCG 332
|||
QY 241 CCCCTCTTCTGGCGAGGAGGATTTCTCCCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 300
|||
Db 333 CCCCTCTTCTGGCGAGGAGGATTTCTCCCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 392
|||
QY 301 GAGCTGGACACCTCCATGGATGGGACTGAGCCCTCAGAAATCCAGTGCACCTCCCTTGGC 360
|||
Db 393 GAGCTGGACACCTCCATGGATGGGACTGAGCCCTCAGAAATCCAGTGCACCTCCCTTGGC 452
|||
QY 361 CTCCAGAAATGAAGTGCACACCCAGCCTGATCCAGTCTTTTAGAAGCTCTGAGCTCCCGG 420
|||
Db 453 CTCCAGAAATGAAGTGCACACCCAGCCTGATCCAGTCTTTTAGAAGCTCTGAGCTCCCGG 512
|||
QY 421 TACTTGGGGACTCTGGCCTGGATGACTTCTTTCTGGACATTGACACATCTGCGGTAGAA 480
|||
Db 513 TACTTGGGGACTCTGGCCTGGATGACTTCTTTCTGGACATTGACACATCTGCGGTAGAA 572
|||
QY 481 AAGGAGCCTGCACGGGCCACACAGAGCCTCCTCAAAACCTCTTCTGTGCCCCAGGTTCT 540
|||
Db 573 AAGGAGCCTGCACGGGCCACACAGAGCCTNCTCAAAACCTCTTCTGTGCCCCAGGTTCT 632
|||
QY 541 TGGGAGTGGAAATGAATCGATCACAATCATTGGAAATCATTTCTGGGTCCTAA 591
|||
Db 633 TGGGAGTGGAAATGAATCGATCACAATCATTGGAAATCATTTCTGGGTCCTAA 683
|||

RESULT 12
BD124690
LOCUS BD124690 825 bp DNA linear PAT 18-SEP-2002
DEFINITION Primer for synthesizing full-length cDNA and use thereof.
ACCESSION BD124690
VERSION BD124690.1 GI:23219635
KEYWORDS JP 2002017375-A/121.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 825)
AUTHORS Ota,T., Nishikawa,T., Isogai,T., Hayashi,K., Ishii,S., Kawai,Y.,
Wakamatsu,A., Sugiyama,T., Nagai,K., Kojima,S., Otsuki,T. and
Koga,H.
TITLE Primer for synthesizing full-length cDNA and use thereof
JOURNAL Patent: JP 2002017375-A 121 22-JAN-2002;
HELIX RESEARCH INSTITUTE
COMMENT OS Homo sapiens (human)
PN JP 2002017375-A/121
PD 22-JAN-2002
PF 07-JUL-2000 JP 2000253172
PI TOSHIO OTA,TETSUO NISHIKAWA,TAKAO ISOGAI,KOJI HAYASHI,SHIZUKO
PI ISHII,
PI YURI KAWAI,AI WAKAMATSU,TOMOYASU SUGIYAMA,KEIICHI NAGAI, PI
SHINICHI KOJIMA,
PI TETSUJI OTSUKI,HISASHI KOGA
PC C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/ PC
10,
PC C12P21/02,C12Q1/68//C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC
Primer for synthesizing full-length cDNA and use thereof FH Key
Location/Qualifiers
FT source 1..825
FT /organism='Homo sapiens (human)'.
FEATURES
source 1..825
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 99.6%; Score 588.4; DB 6; Length 825;
Best Local Similarity 99.7%; Pred. No. 2.6e-119;
Matches 589; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 ATGGAGGGAGGCTTGAAGAGGAACACTCTGATTTGGAAGAGGAGGAGGAGGAGGTGGGAG 60
|||
Db 93 ATGGTGGAGGCTTGAAGAGGAACACTCTGATTTGGAAGAGGAGGAGGAGGAGGTGGGAG 152
|||
QY 61 TGGAGTCCAGCAGGCCTTCAGAGCTACCAGCAAGCCCTGCTCCGCATCTCCCTAGACAAA 120
|||
Db 153 TGGAGTCCAGCAGGCCTTCAGAGCTACCAGCAAGCCCTGCTCCGCATCTCCCTAGACAAA 212
|||
QY 121 GTCCAGCGCAGCCTGGGCCCCCGAGCACCCAGCCCTCCGCAGGCATGTCCTCATCCATAAC 180
|||
Db 213 GTCCAGCGCAGCCTGGGCCCCCGAGCACCCAGCCCTCCGCAGGCATGTCCTCATCCATAAC 272
|||
QY 181 ACCCTCCAACAGCTGCAGGCTGCACCTTCGCCTGGCTCCGCCCTGCCCTGCCCTGCCCTGCCG 240
|||
Db 273 ACCCTCCAACAGCTGCAGGCTGCACCTTCGCCTGGCTCCGCCCTGCCCTGCCCTGCCCTGCCG 332
|||
QY 241 CCCCTCTTCTGGCGAGGAGGATTTCTCCCTGTGAGCCACCATTTGGCTCTATCCTCAGG 300
|||
Db 333 CCCCTCTTCTGGCGAGGAGGATTTCTCCCTGTGAGCCACCATTTGGCTCTATCCTCAGG 392
|||
QY 301 GAGCTGGACACCTCCATGGATGGGACTGAGCCCTCAGAAATCCAGTGCACCTCCCTTGGC 360
|||
Db 393 GAGCTGGACACCTCCATGGATGGGACTGAGCCCTCAGAAATCCAGTGCACCTCCCTTGGC 452
|||
QY 361 CTCCAGAAATGAAGTGCACACCCAGCCTGATCCAGTCTTTTAGAAGCTCTGAGCTCCCGG 420
|||

Db 453 CTCCAGAATGAAGTGCACACCCAGCCTGATCCAGTCTTCTTAGAAGCTCTGAGCTCCCGG 512

Qy 421 TACTTGGGGGACTCTGGCCTGGATGACTTCTTTCTGGACATTGACACATCTGCGGTAGAA 480

Db 513 TACTTGGGGGACTCTGGCCTGGATGACTTCTTTCTGGACATTGACACATCTGCGGTAGAA 572

Qy 481 AAGGAGCCTGCACGGGCCCCCACCAGAGCCTCCTCACAACCTCTTCTGTGCCCCAGGTTCT 540

Db 573 AAGGAGCCTGCACGGGCCCCCACCAGAGCCTNCTCACAACCTCTTCTGTGCCCCAGGTTCT 632

Qy 541 TGGGAGTGAATGAATGGATCACATCATGGAATCATTTCTGGGTCTTAA 591

Db 633 TGGGAGTGAATGAATGGATCACATCATGGAATCATTTCTGGGTCTTAA 683

RESULT 13

BD126646

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

OS

PN

PD

PF

PI

PI

PI

SHINICHI KOJIMA,

PI

PC

C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/ 10,

PC

C12P21/02,C12Q1/68//C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC

Primer for synthesizing full-length cDNA and use thereof FH

Key

Location/Qualifiers

FT

source

1. .825

/organism='Homo sapiens (human)'

Location/Qualifiers

1. .825

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

ORIGIN

Query Match

Best Local Similarity

Matches

589; Conservative

0; Mismatches

2; Indels

0; Gaps

0;

Qy 1 ATGGAGGGAGGCTTTGAAGAGGAAACACTCTGATTTGGAAGAGGAGGAGGAGGAGGTTGGAG 60

Db 93 ATGGTGGGAGGCTTTGAAGAGGAAACACTCTGATTTGGAAGAGGAGGAGGAGGAGGTTGGAG 152

Qy 61 TGGAGTCCAGCAGGCTTCAGAGCTACCAAGCAGCCCTGCTCCGATCTCCCTAGACAAA 120

Db 153 TGGAGTCCAGCAGGCTTCAGAGCTACCAAGCAGCCCTGCTCCGATCTCCCTAGACAAA 212

Qy 121 GTCCAGCGCAGCCTTGGGCCCCCGAGCACCCAGCCTCCGAGGATGTCCTCATCCATAAC 180

Db 213 GTCCAGCGCAGCCTTGGGCCCCCGAGCACCCAGCCTCCGAGGATGTCCTCATCCATAAC 272

Qy 181 ACCCTCCAACAGCTGCAGGCTGCACCTTGCCTCGGCTCCCGCCCCCTGCCCTGCCCTCCCGAG 240

Db 273 ACCCTCCAACAGCTGCAGGCTGCACCTTGCCTCGGCTCCCGCCCCCTGCCCTGCCCTCCCGAG 332

Qy 241 CCCCTCTTCTTGGGGGAGGAGGATTTCTCCCTGTCTAGCCACCATTGGGCTCTATCTCCTCAGG 300

Db 333 CCCCTCTTCTTGGGGGAGGAGGATTTCTCCCTGTCTAGCCACCATTGGGCTCTATCTCCTCAGG 392

Qy 301 GAGCTGGACACCTCCATGGATGGGACTGAGCCCCCTCAGAAATCCAGTGAATCCCTTGGC 360

Db 393 GAGCTGGACACCTCCATGGATGGGACTGAGCCCCCTCAGAAATCCAGTGAATCCCTTGGC 452

Qy 361 CTCCAGAATGAATGAGTCCACCCAGCCTGATCAGTCTTCTTAGAAGCTCTGAGCTCCCGG 420

Db 453 CTCCAGAATGAATGAGTCCACCCAGCCTGATCAGTCTTCTTAGAAGCTCTGAGCTCCCGG 512

Qy 421 TACTTGGGGGACTCTGGCCTGGATGACTTCTTTCTGGACATTGACACATCTGCGGTAGAA 480

Db 513 TACTTGGGGGACTCTGGCCTGGATGACTTCTTTCTGGACATTGACACATCTGCGGTAGAA 572

Qy 481 AAGGAGCCTGCACGGGCCCCCACCAGAGCCTCCTCACAACCTCTTCTGTGCCCCAGGTTCT 540

Db 573 AAGGAGCCTGCACGGGCCCCCACCAGAGCCTNCTCACAACCTCTTCTGTGCCCCAGGTTCT 632

Qy 541 TGGGAGTGAATGAATGGATCACATCATGGAATCATTTCTGGGTCTTAA 591

Db 633 TGGGAGTGAATGAATGGATCACATCATGGAATCATTTCTGGGTCTTAA 683

RESULT 14

AC021625

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

AC021625

Homo sapiens clone RP11-384B6, WORKING DRAFT SEQUENCE, 34 unordered pieces.

AC021625

GI:7331453

HTG; HTGS_PHASE1; HTGS_DRAFT.

Homo sapiens (human)

Homo sapiens

Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 164652)

Birren,B., Linton,L., Nusbaum,C. and Lander,E.

Homo sapiens, clone RP11-384B6

Unpublished

2 (bases 1 to 164652)

Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N., Anderson,S., Baldwin,J., Barna,N., Beckerly,R., Beda,F., Boguslavkiy,L., Boukhgalter,B., Brown,A., Burkett,G., Castle,A., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P., DeArellano,K., Dewar,K., Domino,M., Doyle,M., Fenestor,J., Ferreira,P., FitzHugh,W., Forrest,C., Gage,D., Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L., Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J., Landers,T., Lehoczy,J., Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K., McPheeters,R., Meldrim,J., Meneus,L., Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P., Olivar,T.M., Peterson,K., Pierre,N., Pisani,C., Pollara,V., Raymond,C., Riley,R., Rothman,D., Roy,A., Santos,R., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Tirrell,A., Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.

Direct Submission

Submitted (16-JAN-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

On Mar 28, 2000 this sequence version replaced gi:6705474.

All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)

http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

```
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L5880
Center clone name: 384_E_6
----- Summary Statistics
Sequencing vector: M13; M77815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 145359 bases at least Q40
Consensus quality: 155420 bases at least Q30
Consensus quality: 159351 bases at least Q20
Insert size: 161352; sum-of-contigs
Quality coverage: 4.0 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 34 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 1083: contig of 1083 bp in length
* 1084 1183: gap of 100 bp
* 1184 2479: contig of 1296 bp in length
* 2480 2579: gap of 100 bp
* 2580 4311: contig of 1732 bp in length
* 4312 4411: gap of 100 bp
* 4412 5951: contig of 1540 bp in length
* 5952 6051: gap of 100 bp
* 6052 7761: contig of 1710 bp in length
* 7762 7861: gap of 100 bp
* 7862 9630: contig of 1769 bp in length
* 9631 9730: gap of 100 bp
* 9731 11763: contig of 2033 bp in length
* 11764 11863: gap of 100 bp
* 11864 14294: contig of 2431 bp in length
* 14295 14394: gap of 100 bp
* 14395 16731: contig of 2337 bp in length
* 16732 16831: gap of 100 bp
* 16832 20279: contig of 3448 bp in length
* 20280 20379: gap of 100 bp
* 20380 23751: contig of 3372 bp in length
* 23752 23851: gap of 100 bp
* 23852 26903: contig of 3052 bp in length
* 26904 27003: gap of 100 bp
* 27004 31228: contig of 4225 bp in length
* 31229 31328: gap of 100 bp
* 31329 35336: contig of 4008 bp in length
* 35337 35436: gap of 100 bp
* 35437 38469: contig of 3033 bp in length
* 38470 38569: gap of 100 bp
* 38570 42643: contig of 4074 bp in length
* 42644 42743: gap of 100 bp
* 42744 47087: contig of 4344 bp in length
* 47088 47187: gap of 100 bp
* 47188 51475: contig of 4288 bp in length
* 51476 51575: gap of 100 bp
* 51576 56093: contig of 4518 bp in length
* 56094 56193: gap of 100 bp
* 56194 59387: contig of 3194 bp in length
* 59388 59487: gap of 100 bp
* 59488 63987: contig of 4500 bp in length
* 63988 64087: gap of 100 bp
* 64088 69347: contig of 5260 bp in length
* 69348 69447: gap of 100 bp
* 69448 74268: contig of 4821 bp in length
* 74269 74368: gap of 100 bp
* 74369 79365: contig of 4997 bp in length
* 79366 79465: gap of 100 bp
* 79466 83211: contig of 3746 bp in length
* 83212 83311: gap of 100 bp
```

```
* 83312 88101: contig of 4790 bp in length
* 88102 88201: gap of 100 bp
* 88202 93785: contig of 5584 bp in length
* 93786 93885: gap of 100 bp
* 93886 100407: contig of 6522 bp in length
* 100408 100507: gap of 100 bp
* 100508 108158: contig of 7651 bp in length
* 108159 108258: gap of 100 bp
* 108259 115631: contig of 7373 bp in length
* 115632 115731: gap of 100 bp
* 115732 125488: contig of 9757 bp in length
* 125489 125588: gap of 100 bp
* 125589 136013: contig of 10425 bp in length
* 136014 136113: gap of 100 bp
* 136114 149110: contig of 12997 bp in length
* 149111 149210: gap of 100 bp
* 149211 164652: contig of 15442 bp in length.
*
FEATURES
      source
      1..164652
      /organism="Homo sapiens"
      /mol_type="genomic DNA"
      /db_xref="taxon:9606"
      /clone="RP11-384E6"
      /clone_lib="RPC1-11 Human Male BAC"
      1..1083
      /note="assembly_fragment"
      1184..2479
      /note="assembly_fragment"
      2580..4311
      /note="assembly_fragment"
      4412..5951
      /note="assembly_fragment"
      6052..7761
      /note="assembly_fragment"
      7862..9630
      /note="assembly_fragment"
      9731..11763
      /note="assembly_fragment"
      11864..14294
      /note="assembly_fragment"
      14395..16731
      /note="assembly_fragment"
      16832..20279
      /note="assembly_fragment"
      20380..23751
      /note="assembly_fragment"
      23852..26903
      /note="assembly_fragment"
      27004..31228
      /note="assembly_fragment"
      31329..35336
      /note="assembly_fragment"
      35437..38469
      /note="assembly_fragment"
      38570..42643
      /note="assembly_fragment"
      42744..47087
      /note="assembly_fragment"
      47188..51475
      /note="assembly_fragment"
      51576..56093
      /note="assembly_fragment"
      56194..59387
      /note="assembly_fragment"
      59488..63987
      /note="assembly_fragment"
      64088..69347
      /note="assembly_fragment"
      69448..74268
      /note="assembly_fragment"
      74369..79365
      /note="assembly_fragment"
      79466..83211
      /note="assembly_fragment"
```

```
misc_feature      /note="assembly_fragment"
83312..88101
misc_feature      /note="assembly_fragment"
88202..93785
misc_feature      /note="assembly_fragment"
93886..100407
clone_end:T7
vector_side:right
misc_feature      100508..108158
misc_feature      /note="assembly_fragment"
108259..115631

Query Match      98.9%; Score 584.6; DB 2; Length 164652;
Best Local Similarity 99.3%; Pred. No. 9e-119;
Matches 587; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ATGGAGGGAGGCTTGAAGAGGAAACACTCTGATTTTGGAAAGAGGAGGAGAGGTGGGAG 60
|||||
Db 113509 ATGGTGGGAGGCTTGAAGAGGAAACACTCTGATTTTGGAAAGAGGAGGAGAGGTGGGAG 113568

QY 61 TGGAGTCCAGCAGGCCCTTCAAGAGCTACCAGCAAGCCCTGCTCCGCATCTCCCTAGACAAA 120
|||||
Db 113569 TGGAGTCCAGCAGGCCCTTCAAGAGCTACCAGCAAGCCCTGCTCCGCATCTCCCTAGACAAA 113628

QY 121 GTCCAGCGCAGCCTGGGCCCCCGAGCACCAGCCTCCGAGGCGATGTCTCTATCCATAAC 180
|||||
Db 113629 GTCCAGCGCAGCCTGGGCCCCCGAGCACCAGCCTCCGAGGCGATGTCTCTATCCATAAC 113688

QY 181 ACCCTCCAACAGCTGCAGGCTGCACATTGCGCTGGCTCCGCCCCCTGCCCCCGAG 240
|||||
Db 113689 ACCCTCCAACAGCTGCAGGCTGCACATTGCGCTGGCTCCGCCCCCTGCCCCCGAG 113748

QY 241 CCCCTCTTCTCTGGGCGAGGAGGATTTCTCCCTGTGAGCCACCATTTGGCTCTATCTCAGG 300
|||||
Db 113749 CCCCTCTTCTCTGGGCGAGGAGGATTTCTCCCTGTGAGCCACCATTTGGCTCTATCTCAGG 113808

QY 301 GAGCTGGACACCTCCATGGATGGGACTGAGCCCCCTCAGAATCCAGTGACTCCCTTGGC 360
|||||
Db 113809 GAGCTGGACACCTCCATGGATGGGACTGAGCCCCCTCAGAATCCAGTGACTCCCTTGGC 113868

QY 361 CTCCAGAATGAAGTGCCACCCAGCCTGATCCAGTCTTCTTAGAAGCTCTGAGCTCCCGG 420
|||||
Db 113869 CTCCAGAATGAAGTGCCACCCAGCCTGATCCAGTCTTCTTAGAAGCTCTGAGCTCCCGG 113928

QY 421 TACTTGGGGGACTCTGGCCCTGGATGACTTCTTTCTGGACATTGACACATCTGCGGTAGAA 480
|||||
Db 113929 TACTTGGGGGACTCTGGCCCTGGATGACTTCTTTCTGGACATTGACACATATGCGGTAGAA 113988

QY 481 AAGGAGCCTGCACGGGCCCCCAGAGCCTCCTCAACAACCTTCTGTGCCCCCAGGTTCT 540
|||||
Db 113989 AAGGAGCCTGCACGGGCCCCCAGAGCCTCCTCAACAACCTTCTGTGCCCCCAGGTTCT 114048

QY 541 TGGGAGTGGAAATGAATGGATCACAATCATGGAAATCATTTCTGGGTCTCTAA 591
|||||
Db 114049 TGGGAGTGGAAATGAATGGATCACAATCATGGAAATCATTTCTGGGTCTCTAA 114099
```

```
RESULT 15
CQ463286
LOCUS      CQ463286          506 bp      DNA      linear      PAT 30-JAN-2004
DEFINITION Sequence 6064 from Patent WO0192581.
ACCESSION  CQ463286
VERSION    CQ463286.1  GI:41428905
KEYWORDS
SOURCE     Homo sapiens (human)
ORGANISM  Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1
AUTHORS   Algate,P.A., Harlocker,S.L. and Jones,R.
TITLE     Compositions and methods for the therapy and diagnosis of
ovariancancer
```

```
JOURNAL Patent: WO 0192581-A 6064 06-DEC-2001;
CORIXA CORPORATION (US)
FEATURES Location/Qualifiers
source 1..506
        /organism="Homo sapiens"
        /mol_type="unassigned DNA"
        /db_xref="taxon:9606"
ORIGIN
Query Match      80.0%; Score 473; DB 6; Length 506;
Best Local Similarity 98.6%; Pred. No. 7.4e-94;
Matches 488; Conservative 0; Mismatches 5; Indels 2; Gaps 1;

QY 60 GTGGAGTCCAGCAGGCGCTTTCAGAGCTACCAGCAAGCCCTGCTCCGCATCTCCCTAGACAA 119
|||||
Db 14 GCGGAGTCCAGCAGGCGCTTTCAGAGCTACCAGCAAGCCCTGCTCCGCATCTCCCTAGACAA 73

QY 120 AGTCCAGCGCAGCCTGGGCCCCCGAGCACCAGCCTCCGCGAGGCATGTCTCATCCATAA 179
|||||
Db 74 AGTCCAGCGC--CCTGGGCCCCCGAGCACCAGCCTCCGCGAGCATGTCTCATCCATAA 131

QY 180 CACCCCTCCAACAGCTGCAGGCTGCACATTGCGCTGGCTCCGCCCCCTGCCCTGCCCCCGA 239
|||||
Db 132 CACCCCTCCAACAGCTGCAGGATGCACATTGCGCTGGCTCCGCCCCCTGCCCTGCCCCCGA 191

QY 240 GCCCCTCTTCTTGGCGGAGGAGGATTTCTCCCTGTGAGCCACCATTTGGCTCTATCTCAG 299
|||||
Db 192 GCCCCTCTTCTTGGCGGAGGAGGATTTCTCCCTGTGAGCCACCATTTGGCTCTATCTCAG 251

QY 300 GGAGCTGGACACCTCCATGGATGGGACTGAGCCCCCTCAGAATCCAGTGACTCCCTTGG 359
|||||
Db 252 GGAGCTGGACACCTCCATGGATGGGACTGAGCCCCCTCAGAATCCAGTGACTCCCTTGG 311

QY 360 CCTCCAGAATGAAGTGCCACCCAGCCTGATCCAGTCTTCTTAGAAGCTCTGAGCTCCCG 419
|||||
Db 312 CCTCCAGAATGAAGTGCCACCCAGCCTGATCCAGTCTTCTTAGAAGCTCTGAGCTCCCG 371

QY 420 GTACTTGGGGGACTCTGGCCCTGGATGACTTCTTTCTTGGACATTGACACATCTGCGGTAGA 479
|||||
Db 372 GTACTTGGGGGACTCTGGCCCTGGATGACTTCTTTCTTGGACATTGACACATCTGCGGTAGA 431

QY 480 AAAGGAGCCTGCACGGGCCCCCAGAGCCTCCTCACAACCTCTTCTGTGCCCCAGGTTTC 539
|||||
Db 432 AAAGGAGCCTGCACGGGCCCCCAGAGCCTCCTCACAACCTCTTCTGTGCCCCAGGTTTC 491

QY 540 TTGGGAGTGGAAATGA 554
|||||
Db 492 TTGGGAGTGGAAATGA 506
```

Search completed: April 26, 2005, 09:13:33
Job time : 2755 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model
Run on: April 26, 2005, 03:13:57 ; Search time 425 Seconds
(without alignments)
8231.923 Million cell updates/sec

Title: US-10-069-386A-1
Perfect score: 591
Sequence: 1 atggaggaggcttgaagag.....aatcattctgggtcctaa 591

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues 8780412
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002as:*
7: Geneseqn2002bs:*
8: Geneseqn2003as:*
9: Geneseqn2003bs:*
10: Geneseqn2003cs:*
11: Geneseqn2003ds:*
12: Geneseqn2004as:*
13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query %			DB	ID	Description
	Score	Match	Length			
1	591	100.0	591	4	AAF28052	Aaf28052 Replicati
2	589.4	99.7	1325	4	AAK94884	Aak94884 Human ful
3	589.4	99.7	1325	12	ADL32051	Adl32051 Full leng
4	589.4	99.7	1515	13	ADQ85787	Adq85787 Human tum
5	589.4	99.7	1542	10	ADG33127	Adg33127 Human DNA
6	588.4	99.6	825	4	AAK91661	Aak91661 Human cDN
7	588.4	99.6	825	4	AAK93617	Aak93617 Human cDN
8	588.4	99.6	825	12	ADL30044	Adl30044 3' end of
9	588.4	99.6	825	12	ADL28088	Adl28088 5' end of
10	578.4	97.9	598	12	ACH91703	Ach91703 Human gen
11	576.2	97.5	1596	4	AAH34147	Aah34147 Human col
12	473	80.0	506	6	ABL83086	Ab183086 Human ova
13	186.6	31.6	205	6	ABL83261	Ab183261 Human ova
14	166	28.1	526	12	ACH77983	Ach77983 Human gen
15	47.4	8.0	2082	6	AAL41497	Aal41497 Mouse hae
c 16	46.4	7.9	51552	6	AAS96607	Aas96607 DNA encod
17	42.6	7.2	1603	8	ACC51095	Acc51095 Human Sho
18	42.4	7.2	511	12	ACH91786	Ach91786 Human gen
19	42.4	7.2	600	6	ABQ52497	Abq52497 Oligonucl
c 20	42.4	7.2	600	6	ABQ52496	Abq52496 Oligonucl

c	21	42.4	7.2	1850	8	ADA70325	Ada70325 Rice gene
c	22	41.8	7.1	32329	12	ADO51695	Ado51695 Streptomy
	23	41.4	7.0	825	12	ADO00165	Ado00165 Novel hum
	24	41.4	7.0	825	12	ADN98596	Adn98596 Novel hum
c	25	41	6.9	1925	2	AAX90924	Aax90924 Epstein B
c	26	41	6.9	1926	3	AAA50254	Aaa50254 Epstein B
c	27	41	6.9	1926	4	AAF82902	Aaf82902 EBV tethe
c	28	41	6.9	1926	10	ADK65580	Adk65580 Human her
c	29	41	6.9	2580	3	AAA75454	Aaa75454 Nucleotid
c	30	41	6.9	2580	6	AAI64275	Aai64275 Epstein-B
	31	41	6.9	5452	2	AAX90923	Aax90923 Anti-sens
	32	41	6.9	8705	2	AAZ23778	Aaz23778 Vector ps
	33	41	6.9	8705	12	ADM10659	Adm10659 Expressio
c	34	41	6.9	9482	12	ADP64415	Adp64415 Vector pC
	35	41	6.9	9600	2	AAV21683	Aav21683 Vector pl
	36	41	6.9	10285	6	ABS71027	Abx71027 pCEP-Xa-F
c	37	41	6.9	10285	6	ABS66453	Abx66453 Plasmid p
	38	41	6.9	10330	12	ADL67154	Adl67154 Plasmid p
	39	41	6.9	10380	2	AAZ22248	Aaz22248 Nucleotid
c	40	41	6.9	10477	12	ADL67152	Adl67152 Plasmid p
	41	41	6.9	10516	12	ADL67150	Adl67150 Plasmid p
	42	41	6.9	10561	12	ADL67148	Adl67148 Plasmid p
c	43	41	6.9	10596	2	AAQ51731	Aaq51731 Plasmid p
	44	41	6.9	10596	2	AAX15650	Aax15650 Nucleotid
	45	41	6.9	10596	2	AAT40348	Aat40348 Plasmid p

ALIGNMENTS

RESULT 1
AAF28052
ID AAF28052 standard; DNA; 591 BP.
XX
AC AAF28052;
XX
DT 23-MAY-2001 (first entry)
XX
DE Replication protein A binding transcriptional activator 1 RBT1 gene.
XX
KW RBT1; replication protein A binding transcriptional activator 1; RPA32;
KW gene therapy; apoptosis; cancer; leukaemia; ds.
XX
OS Unidentified.
XX
FH Key Location/Qualifiers
FT CDS 1..591
FT /*tag= a
FT /product= "RBT1"

WO200114546-A2.
XX
PD 01-MAR-2001.
XX
PF 17-AUG-2000; 2000WO-CA000948.
XX
PR 19-AUG-1999; 99US-0149472P.
XX
PA (TRAN-) CENT TRANSLATIONAL RES IN CANCER.
XX
PI Alaoui-Jamali MA, Cho JM;
XX
DR WPI; 2001-218447/22.
DR P-PSDB; AAB35402.
XX
PT Novel replication protein A binding transcriptional activator 1 gene,
PT useful for treating neoplastic disorders such as cancer and in gene
PT therapy.
XX
PS Claim 1; Fig 1; 16pp; English.
XX
CC The present invention provides the protein and coding sequences of the
CC replication protein A binding transcriptional activator 1 (RBT1). The

CC protein is capable of inducing apoptosis. The sequences are useful in the
CC gene therapy and other methods of treatment of cancer, including
CC leukaemias. The present sequence is the RBT1 coding sequence

XX
SQ Sequence 591 BP; 123 A; 194 C; 156 G; 118 T; 0 U; 0 Other;
Query Match 100.0%; Score 591; DB 4; Length 591;
Best Local Similarity 100.0%; Pred. No. 4.3e-135;
Matches 591; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGAGGGAGGCTTGAAGAGAAACACTCTGATTTGGAAGAGGAGGAGGAGGAGGTGGGAG 60
Db |||||
QY 61 TGGAGTCCAGCAGGCCCTTGAGAGGAAACACTCTGATTTGGAAGAGGAGGAGGAGGTGGGAG 60
Db |||||
QY 61 TGGAGTCCAGCAGGCCCTTGAGAGGAAACACTCTGATTTGGAAGAGGAGGAGGAGGTGGGAG 60
Db |||||
QY 121 GTCCAGCGCAGCCTGGGCCCCCGAGCACCCAGCCTCCGCGAGGCATGTCTCATCCATAAC 180
Db |||||
QY 121 GTCCAGCGCAGCCTGGGCCCCCGAGCACCCAGCCTCCGCGAGGCATGTCTCATCCATAAC 180
Db |||||
QY 181 ACCCTCCAACAGCTGCAGGTGCACCTTGCCTGGCTCCGCCCCCTGCCCTGCCCCCGGAG 240
Db |||||
QY 181 ACCCTCCAACAGCTGCAGGTGCACCTTGCCTGGCTCCGCCCCCTGCCCTGCCCCCGGAG 240
Db |||||
QY 241 CCCCTCTTCTGGGCGAGGAGGATTTCTCCCTGTCCAGCCACCATTGGGCTCTATCCTCAGG 300
Db |||||
QY 241 CCCCTCTTCTGGGCGAGGAGGATTTCTCCCTGTCCAGCCACCATTGGGCTCTATCCTCAGG 300
Db |||||
QY 301 GAGCTGGACACCTCCATGGATGGGACTGAGCCCCCTCAGAATCCAGTGACTCCCTTGGC 360
Db |||||
QY 301 GAGCTGGACACCTCCATGGATGGGACTGAGCCCCCTCAGAATCCAGTGACTCCCTTGGC 360
Db |||||
QY 361 CTCCAGAAATGAAGTGCCACCCAGCCTGATCCAGTCTTTAGAAAGCTCTGAGTCCCGG 420
Db |||||
QY 361 CTCCAGAAATGAAGTGCCACCCAGCCTGATCCAGTCTTTAGAAAGCTCTGAGTCCCGG 420
Db |||||
QY 421 TACTTGGGGACTCTGGCCTGGATGACTTCTTTCTGGACATTGACACATCTGCGGTAGAA 480
Db |||||
QY 421 TACTTGGGGACTCTGGCCTGGATGACTTCTTTCTGGACATTGACACATCTGCGGTAGAA 480
Db |||||
QY 481 AAGGAGCCTGCACGGGCCCCACAGAGCCTCCTCACAACCTTCTTGCCCCCAGGTTCT 540
Db |||||
QY 481 AAGGAGCCTGCACGGGCCCCACAGAGCCTCCTCACAACCTTCTTGCCCCCAGGTTCT 540
Db |||||
QY 541 TGGGAGTGGAAATGAACATCATGGAATCAATTTCTGGGGTCTCTAA 591
Db |||||
QY 541 TGGGAGTGGAAATGAACATCATGGAATCAATTTCTGGGGTCTCTAA 591
Db |||||

RESULT 2
AAK94884
ID AAK94884 standard; cDNA; 1325 BP.
AC AAK94884;
XX
DT 06-NOV-2001 (first entry)
DE Human full-length cDNA, SEQ ID NO: 4084.
XX

KW Human; full length cDNA; cDNA synthesis; oligo-capping; ss.
XX Homo sapiens.
XX EP1130094-A2.
PD 05-SEP-2001.
XX
PF 07-JUL-2000; 2000EP-00114089.
XX
PR 08-JUL-1999; 99JP-00194486.
PR 11-JAN-2000; 2000JP-00118774.

PR 02-MAY-2000; 2000JP-00183765.
XX (HELI-) HELIX RES INST.
PA
XX
PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX
DR WPI; 2001-524255/58.
DR P-PSDB; AAM93922.
XX
PT 830 Primers useful for synthesizing full length cDNA clones and their use
PT in genetic manipulation.
XX
PS Claim 8; SEQ ID NO 4084; 1380pp + Sequence Listing; English.
XX
CC The invention relates to primers for synthesising full length cDNA
CC clones. 830 cDNA molecules encoding a human protein have been isolated
CC and nucleotide sequences of 5'- and 3'-ends of the cDNA molecules have
CC been determined. Primers for synthesising the full length cDNA are useful
CC for clarifying the function of the protein encoded by the cDNA. The full
CC length clones were obtained by construction of full length enriched cDNA
CC libraries that were synthesised by the oligo-capping method. The primers
CC enable the production of the full length cDNA easily without any special
CC methods. The present sequence is a full length human cDNA of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in CD-ROM format directly
CC from EPO
XX

SQ Sequence 1325 BP; 263 A; 371 C; 359 G; 332 T; 0 U; 0 Other;
Query Match 99.7%; Score 589.4; DB 4; Length 1325;
Best Local Similarity 99.8%; Pred. No. 1.3e-134;
Matches 590; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGAGGGAGGCTTGAAGAGGAAACACTCTGATTTGGAAGAGGAGGAGGAGGTGGGAG 60
Db |||||
QY 93 ATGGTGGGAGGCTTGAAGAGGAAACACTCTGATTTGGAAGAGGAGGAGGAGGTGGGAG 152
Db |||||
QY 61 TGGAGTCCAGCAGGCCCTTCCAGAGCTACCAGCAAGCCCTGCTCCGCATCTCCCTAGACAAA 120
Db |||||
QY 153 TGGAGTCCAGCAGGCCCTTCCAGAGCTACCAGCAAGCCCTGCTCCGCATCTCCCTAGACAAA 212
Db |||||
QY 121 GTCCAGCGCAGCCTGGGCCCCCGAGCACCCAGCCTCCGCGAGGCATGTCTCATCCATAAC 180
Db |||||
QY 213 GTCCAGCGCAGCCTGGGCCCCCGAGCACCCAGCCTCCGCGAGGCATGTCTCATCCATAAC 272
Db |||||
QY 181 ACCCTCCAACAGCTGCAGGCTGCACCTTGCCTTGGCTCCCGCCCCCTGCCCTGCCCCCGGAG 240
Db |||||
QY 273 ACCCTCCAACAGCTGCAGGCTGCACCTTGCCTTGGCTCCCGCCCCCTGCCCTGCCCCCGGAG 332
Db |||||
QY 241 CCCCTCTTCTGGGCGAGGAGGATTTCTCCCTGTCCAGCCACCATTGGGCTCTATCCTCAGG 300
Db |||||
QY 333 CCCCTCTTCTGGGCGAGGAGGATTTCTCCCTGTCCAGCCACCATTGGGCTCTATCCTCAGG 392
Db |||||
QY 301 GAGCTGGACACCTCCATGGATGGGACTGAGCCCCCTCAGAAATCCAGTGACTCCCTTGGC 360
Db |||||
QY 393 GAGCTGGACACCTCCATGGATGGGACTGAGCCCCCTCAGAAATCCAGTGACTCCCTTGGC 452
Db |||||
QY 361 CTCCAGAAATGAAGTGCCACCCAGCCTGATCCAGTCTTTAGAAAGCTCTGAGTCCCGG 420
Db |||||
QY 453 CTCCAGAAATGAAGTGCCACCCAGCCTGATCCAGTCTTTAGAAAGCTCTGAGTCCCGG 512
Db |||||
QY 421 TACTTGGGGACTCTGGCCTGGATGACTTCTTTCTGGACATTGACACATCTGCGGTAGAA 480
Db |||||
QY 513 TACTTGGGGACTCTGGCCTGGATGACTTCTTTCTGGACATTGACACATCTGCGGTAGAA 572
Db |||||
QY 481 AAGGAGCCTGCACGGGCCCCACAGAGCCTCCTCACAACCTTCTTGCCCCCAGGTTCT 540
Db |||||
QY 573 AAGGAGCCTGCACGGGCCCCACAGAGCCTCCTCACAACCTTCTTGCCCCCAGGTTCT 632
Db |||||
QY 541 TGGGAGTGGAAATGAACATCATGGAATCAATTTCTGGGGTCTCTAA 591
Db |||||
QY 633 TGGGAGTGGAAATGAACATCATGGAATCAATTTCTGGGGTCTCTAA 683
Db |||||

RESULT 3
ADL32051
ID ADL32051 standard; cDNA; 1325 BP.
XX
AC ADL32051;
XX
DT 20-MAY-2004 (first entry)
XX
DE Full length human cDNA clone SeqID 4084.
XX
KW human; medicine; signal transduction; glycoprotein; transcription;
KW oligo-capping method; ss; gene.
XX
OS Homo sapiens.
XX
PN EP1396543-A2.
XX
PD 10-MAR-2004.
XX
PF 07-JUL-2000; 2003EP-00025638.
XX
PR 08-JUL-1999; 99JP-00194486.
PR 11-JAN-2000; 2000JP-00118774.
PR 02-MAY-2000; 2000JP-00183865.
PR 07-JUL-2000; 2000EP-00114089.
XX
PA (REAS-) RES ASSOC BIOTECHNOLOGY.
XX
PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX
DR WPI; 2004-204755/20.
DR P-PSDB; ADL32052.
XX
XX New oligonucleotide primers (830 cDNAs) useful for synthesizing full
length human cDNAs.
XX
PS Example 1; SEQ ID NO 4084; 1340pp; English.
XX
CC This invention relates to a novel primers useful for synthesising full
length cDNA molecules that encode human proteins. Specifically, it refers
to secretory or membrane proteins that are potential therapeutic agents/
target molecules in the field of medicine, and in particular genes
encoding proteins that are associated with signal transduction,
glycoproteins and transcription. The present invention describes a method
for efficiently cloning a full length human cDNA from both the 5' and 3'
ends using the oligo-capping method. This polynucleotide sequence is a
full length human cDNA clone of the invention.
XX
SQ Sequence 1325 BP; 263 A; 371 C; 359 G; 332 T; 0 U; 0 Other;

Query Match 99.7%; Score 589.4; DB 12; Length 1325;
Best Local Similarity 99.8%; Pred. No. 1.3e-134;
Matches 590; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGAGGGAGGCTTGAAGAGGAAACACTCTGATTTGGAAGAGGAGGAGGAGGTGGGAG 60
Db |||||
QY 61 TGGAGTCCAGCAGGCTTCAGAGCTACCAAGCAAGCCCTGCTCCGCATCTCCCTAGACAAA 120
Db |||||
QY 153 TGGAGTCCAGCAGGCTTCAGAGCTACCAAGCAAGCCCTGCTCCGCATCTCCCTAGACAAA 212
Db |||||
QY 121 GTCCAGCGCAGCCTGGGCCCCCGAGCACCAGCCCTCCGAGGCATGTCCTCATCCATAAC 180
Db |||||
QY 213 GTCCAGCGCAGCCTGGGCCCCCGAGCACCAGCCCTCCGAGGCATGTCCTCATCCATAAC 272
QY 181 ACCCTCCAACAGCTGAGGCTGCACCTTGCCTGGCTCCGCCCCCTGCCCTGCCCTCCGAG 240
Db |||||
QY 273 ACCCTCCAACAGCTGAGGCTGCACCTTGCCTGGCTCCGCCCCCTGCCCTGCCCTCCGAG 332
Db |||||

QY 241 CCCCTCTTCTGGCGAGGAGGATTTCTCCCTGTGAGCCACCATTTGGCTCTATCCTCAGG 300
Db |||||
QY 333 CCCCTCTTCTGGCGAGGAGGATTTCTCCCTGTGAGCCACCATTTGGCTCTATCCTCAGG 392
Db |||||
QY 301 GAGCTGGACACCTCCATGGATGGGACTGAGCCCTCAGAAATCCAGTACTCCCTTGGC 360
Db |||||
QY 393 GAGCTGGACACCTCCATGGATGGGACTGAGCCCTCAGAAATCCAGTACTCCCTTGGC 452
QY 361 CTCCAGAATGAAGTGCACCCAGCCCTGATCCAGTCTTCTTAGAAGCTCTGAGCTCCCGG 420
Db |||||
QY 453 CTCCAGAATGAAGTGCACCCAGCCCTGATCCAGTCTTCTTAGAAGCTCTGAGCTCCCGG 512
QY 421 TACTTGGGGGACTCTGGCCCTGGATGACTTCTTCTTGGACATTGACACATCTGCGGTAGAA 480
Db |||||
QY 513 TACTTGGGGGACTCTGGCCCTGGATGACTTCTTCTTGGACATTGACACATCTGCGGTAGAA 572
QY 481 AAGGAGCCTGCACGGGCCCCCACCAGAGCCTCCTCACAACCTTCTTGTGCCCCCAGGTTCT 540
Db |||||
QY 573 AAGGAGCCTGCACGGGCCCCCACCAGAGCCTCCTCACAACCTTCTTGTGCCCCCAGGTTCT 632
QY 541 TGGGAGTGAATGAAGTGAATGATCAGATCATCATGGAATCATTTCTGGGGTCTCTAA 591
Db |||||
QY 633 TGGGAGTGAATGAAGTGAATGATCAGATCATCATGGAATCATTTCTGGGGTCTCTAA 683

RESULT 4
ADQ85787
ID ADQ85787 standard; cDNA; 1515 BP.
XX
AC ADQ85787;
XX
DT 07-OCT-2004 (first entry)
XX
DE Human tumour-associated antigenic target (TAT) cDNA sequence #2601.
XX
KW human; tumour-associated antigenic target; TAT; cytostatic; gene therapy;
KW cancer; cell proliferative disorder; gene; ss.
XX
OS Homo sapiens.
XX
PN WO2004060270-A2.
XX
PD 22-JUL-2004.
XX
PF 15-OCT-2003; 2003WO-US029126.
XX
PR 18-OCT-2002; 2002US-0418988P.
PA (GETH) GENENTECH INC.
PA (WUTD/) WU T D.
PA (ZHOU/) ZHOU Y.
XX
PI Wu TD, Zhou Y;
XX
DR WPI; 2004-534300/51.
XX
PT New nucleic acid molecule and encoded polypeptide, for diagnosing,
preventing or treating cell proliferative disorders such as cancer.
XX
PS Claim 1; SEQ ID NO 2601; 5504pp; English.
XX

The present invention describes an isolated tumour-associated antigenic target (TAT) nucleic acid comprising: (a) any of 4622 nucleotide sequences (see SEQ ID NO:1 to 4622); (b) the full-length coding region of (a); (c) the complement of (a) or (b); (d) a sequence that has 80% sequence identity to (a)-(c); or (e) a sequence that hybridises to (a)-(c). Also described: (1) an expression vector comprising the above nucleic acid; (2) a host cell comprising the above expression vector; (3) a process for producing a polypeptide; (4) an isolated polypeptide comprising: (a) an amino acid sequence encoded by any of the above nucleotide sequences; (b) an amino acid sequence encoded by the full-length coding region of the above nucleotide sequences; or (c) a sequence having at least 80% identical to (a) or (b); (5) a chimeric polypeptide

Db 412 GTCCAGCGCAGCCTGGGCCCCGAGCACCAGCCCTCCGAGGCATGTCTCATCATCAAC 471
QY 181 ACCCTCCAACAGCTGCAGGTGCACCTTGGCTGGCTCCGCCCCCTGCCCTGCCCGGAG 240
Db 472 ACCCTCCAACAGCTGCAGGTGCACCTTGGCTGGCTCCGCCCCCTGCCCTGCCCGGAG 531
QY 241 CCCCTCTTCTGGGCGAGGAGGATTCTCCCTGTGAGCCACCATTGGCTCTATCTCAGG 300
Db 532 CCCCTCTTCTGGGCGAGGAGGATTCTCCCTGTGAGCCACCATTGGCTCTATCTCAGG 591
QY 301 GAGCTGGACACCTCCATGATGGGACTGAGCCCCCTCAGAAATCCAGTACTCCCTTGGC 360
Db 592 GAGCTGGACACCTCCATGATGGGACTGAGCCCCCTCAGAAATCCAGTACTCCCTTGGC 651
QY 361 CTCCAGAAATGAAGTGCCACCCAGCCTGATCCAGTCTTCTTAGAAGCTCTGAGTCCCGG 420
Db 652 CTCCAGAAATGAAGTGCCACCCAGCCTGATCCAGTCTTCTTAGAAGCTCTGAGTCCCGG 711
QY 421 TACTTGGGGACTCTGGCCTGGATGACTTCTTCTTGGACATTGACACATCTGCGGTAGAA 480
Db 712 TACTTGGGGACTCTGGCCTGGATGACTTCTTCTTGGACATTGACACATCTGCGGTAGAA 771
QY 481 AAGGAGCCTGCACGGGCCCCCAGAGCCTCCTCACAACCTTCTTGTGCCCCAGGTTCT 540
Db 772 AAGGAGCCTGCACGGGCCCCCAGAGCCTCCTCACAACCTTCTTGTGCCCCAGGTTCT 831
QY 541 TGGGAGTGGAAATGAATGATGATCAGATCATGGAATCATTTGGGGTCTCTAA 591
Db 832 TGGGAGTGGAAATGAATGATGATCAGATCATGGAATCATTTGGGGTCTCTAA 882

RESULT 6

AAK91661
ID AAK91661 standard; cDNA; 825 BP.
XX
AC AAK91661;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human cDNA 5'-end sequence, SEQ ID NO: 121.
XX
KW Human; full length cDNA; cDNA synthesis; oligo-capping; ss.
XX
OS Homo sapiens.
XX
PN EP1130094-A2.
XX
PD 05-SEP-2001.
XX
PF 07-JUL-2000; 2000EP-00114089.
XX
PR 08-JUL-1999; 99JP-00194486.
PR 11-JAN-2000; 2000JP-00118774.
PR 02-MAY-2000; 2000JP-00183765.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX
DR WPI; 2001-524255/58.
XX
PT 830 Primers useful for synthesizing full length cDNA clones and their use
PT in genetic manipulation.
XX
PS Claim 2; SEQ ID NO 121; 1380pp + Sequence Listing; English.
XX

CC The invention relates to primers for synthesising full length cDNA
CC clones. 830 cDNA molecules encoding a human protein have been isolated
CC and nucleotide sequences of 5'- and 3'-ends of the cDNA molecules have
CC been determined. Primers for synthesising the full length cDNA are useful
CC for clarifying the function of the protein encoded by the cDNA. The full
CC length clones were obtained by construction of full length enriched cDNA

CC libraries that were synthesised by the oligo-capping method. The primers
CC enable the production of the full length cDNA easily without any special
CC methods. The present sequence is the nucleotide sequence of the 5'-end of
CC a cDNA provided in the invention. Note: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in CD-
CC ROM format directly from EPO

SQ Sequence 825 BP; 160 A; 247 C; 230 G; 183 T; 0 U; 5 Other;
Query Match 99.6%; Score 588.4; DB 4; Length 825;
Best Local Similarity 99.7%; Pred. No. 2e-134;
Matches 589; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGGAGGGAGGCTTGAAGAGGAAACACTCTGTATTTGGAAGAGGAGGAGGAGGTGGGAG 60
Db 93 ATGGTGGGAGGCTTGAAGAGGAAACACTCTGTATTTGGAAGAGGAGGAGGAGGTGGGAG 152
QY 61 TGGAGTCCAGCAGGCTTTCAGAGCTACCCAGCAAGCCCTGCTCCGATCTCCCTAGACAAA 120
Db 153 TGGAGTCCAGCAGGCTTTCAGAGCTACCCAGCAAGCCCTGCTCCGATCTCCCTAGACAAA 212
QY 121 GTCCAGCGCAGCCTGGGCCCCCGAGCACCCAGCTCCGAGGCATGTCTCATCCATAAC 180
Db 213 GTCCAGCGCAGCCTGGGCCCCCGAGCACCCAGCTCCGAGGCATGTCTCATCCATAAC 272
QY 181 ACCCTCCAACAGCTGCAGGCTGCACCTTCCCTTGGCTCCGCCCCCTGCCCGGAG 240
Db 273 ACCCTCCAACAGCTGCAGGCTGCACCTTCCCTTGGCTCCGCCCCCTGCCCGGAG 332
QY 241 CCCCTCTTCTTGGGCGAGGAGGATTTCTCCCTGTGAGCCACCATTGGCTCTATCTCAGG 300
Db 333 CCCCTCTTCTTGGGCGAGGAGGATTTCTCCCTGTGAGCCACCATTGGCTCTATCTCAGG 392
QY 301 GAGCTGGACACCTTCCATGGATGGGACTGAGCCCCCTCAGAAATCCAGTACTCCCTTGGC 360
Db 393 GAGCTGGACACCTTCCATGGATGGGACTGAGCCCCCTCAGAAATCCAGTACTCCCTTGGC 452
QY 361 CTCCAGAAATGAATGAGTGCACCCAGCCTGATCCAGTCTTTTAGAAGCTCTGAGTCCCGG 420
Db 453 CTCCAGAAATGAATGAGTGCACCCAGCCTGATCCAGTCTTTTAGAAGCTCTGAGTCCCGG 512
QY 421 TACTTGGGGACTCTGGCCTGGATGACTTCTTCTTGGACATTGACACATCTGCGGTAGAA 480
Db 513 TACTTGGGGACTCTGGCCTGGATGACTTCTTCTTGGACATTGACACATCTGCGGTAGAA 572
QY 481 AAGGAGCCTGCACGGGCCCCCAGAGCCTCCTCACAACCTTCTTGTGCCCCAGGTTCT 540
Db 573 AAGGAGCCTGCACGGGCCCCCAGAGCCTCCTCACAACCTTCTTGTGCCCCAGGTTCT 632
QY 541 TGGGAGTGGAAATGAATGAGTGGATCAGATCATGGAATCATTTGGGGTCTCTAA 591
Db 633 TGGGAGTGGAAATGAATGAGTGGATCAGATCATGGAATCATTTGGGGTCTCTAA 683

RESULT 7

AAK93617
ID AAK93617 standard; cDNA; 825 BP.
XX
AC AAK93617;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human cDNA clone representative sequence, SEQ ID NO: 2077.
XX
KW Human; full length cDNA; cDNA synthesis; oligo-capping; ss.
XX
OS Homo sapiens.
XX
PN EP1130094-A2.
XX
PD 05-SEP-2001.
XX
PF 07-JUL-2000; 2000EP-00114089.

XX 08-JUL-1999; 99JP-00194486.
PR 11-JAN-2000; 2000JP-00118774.
PR 02-MAY-2000; 2000JP-00183765.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX
DR WPI; 2001-524255/58.
XX
XX 830 Primers useful for synthesizing full length cDNA clones and their use
PT in genetic manipulation.
XX
XX Example 11; SEQ ID NO 2077; 1380pp + Sequence Listing; English.
XX
CC The invention relates to primers for synthesising full length cDNA
CC clones. 830 cDNA molecules encoding a human protein have been isolated
CC and nucleotide sequences of 5'- and 3'-ends of the cDNA molecules have
CC been determined. Primers for synthesising the full length cDNA are useful
CC for clarifying the function of the protein encoded by the cDNA. The full
CC length clones were obtained by construction of full length enriched cDNA
CC libraries that were synthesised by the oligo-capping method. The primers
CC enable the production of the full length cDNA easily without any special
CC methods. The present sequence was used as the representative sequence
CC from a human clone which was used in homology searches to identify the
CC clone. Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in CD-ROM format directly from
CC EPO
XX
SQ Sequence 825 BP; 160 A; 247 C; 230 G; 183 T; 0 U; 5 Other;

Query Match 99.6%; Score 588.4; DB 4; Length 825;
Best Local Similarity 99.7%; Pred. No. 2e-134;
Matches 589; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGGAGGGAGGCTTGAAGAGGAAACACTCTGATTTGGAAGAGGAGGAGGAGGTGGGAG 60
Db |||||
QY 93 ATGGTGGGAGGCTTGAAGAGGAAACACTCTGATTTGGAAGAGGAGGAGGAGGTGGGAG 152
Db |||||
QY 61 TGGAGTCCAGCAGGCCCTTCAGAGCTACCAAGCAAGCCCTGCTCCGCATCTCCCTAGACAAA 120
Db |||||
QY 153 TGGAGTCCAGCAGGCCCTTCAGAGCTACCAAGCAAGCCCTGCTCCGCATCTCCCTAGACAAA 212
Db |||||
QY 121 GTCCAGCGCAGCCTGGGCCCCCGAGCACCCAGCCTCCGAGGCATGTCCTCATCCATAAC 180
Db |||||
QY 213 GTCCAGCGCAGCCTGGGCCCCCGAGCACCCAGCCTCCGAGGCATGTCCTCATCCATAAC 272
Db |||||
QY 181 ACCCTCCAACAGCTGCAGGTGCACCTTGGCTGGCTCCGCCCTGCCCTGCCCTGCCCTGCCGAG 240
Db |||||
QY 273 ACCCTCCAACAGCTGCAGGTGCACCTTGGCTGGCTCCGCCCTGCCCTGCCCTGCCCTGCCGAG 332
Db |||||
QY 241 CCCCTCTTCTGGGCGGAGGAGGATTTCTCCCTGTTCAGCCACCATTGGCTCTATCCTCAGG 300
Db |||||
QY 333 CCCCTCTTCTGGGCGGAGGAGGATTTCTCCCTGTTCAGCCACCATTGGCTCTATCCTCAGG 392
Db |||||
QY 301 GAGCTGGACACCTCCATGGATGGGACTGAGCCCCCTCAGAAATCCAGTGACTCCCTTGGC 360
Db |||||
QY 393 GAGCTGGACACCTCCATGGATGGGACTGAGCCCCCTCAGAAATCCAGTGACTCCCTTGGC 452
Db |||||
QY 361 CTCCAGAAATGAAGTGCCACCCAGCCTGATCCAGTCTTCTTAGAAGCTCTGAGCTCCCGG 420
Db |||||
QY 453 CTCCAGAAATGAAGTGCCACCCAGCCTGATCCAGTCTTCTTAGAAGCTCTGAGCTCCCGG 512
Db |||||
QY 421 TACTTGGGGACTCTGGCCTGGATGACTTCTTTCTGGACATTTGACACATCTGCGGTAGAA 480
Db |||||
QY 513 TACTTGGGGACTCTGGCCTGGATGACTTCTTTCTGGACATTTGACACATCTGCGGTAGAA 572
Db |||||
QY 481 AAGGAGCCTGCACGGGCCCCAGAGCCCTCCCTCAAAACCTCTTCTGTGCCCCAGGTTCT 540
Db |||||
QY 573 AAGGAGCCTGCACGGGCCCCAGAGCCCTNCTCAAAACCTCTTCTGTGCCCCAGGTTCT 632
Db |||||

QY 541 TGGGAGTGGATGAAGTGGATCACATCATGGAATCATTTCTGGGGTCTTAA 591
Db |||||
QY 633 TGGGAGTGGATGAAGTGGATCACATCATGGAATCATTTCTGGGGTCTTAA 683
Db |||||
RESULT 8
ADL30044
ID ADL30044 standard; cDNA; 825 BP.
XX
AC ADL30044;
XX
DT 20-MAY-2004 (first entry)
XX
DE 3' end of a representative human cDNA cluster SeqID 2077.
XX
KW human; medicine; signal transduction; glycoprotein; transcription;
KW oligo-capping method; ss.
XX
OS Homo sapiens.
XX
PN EPI396543-A2.
XX
PD 10-MAR-2004.
XX
XX 07-JUL-2000; 2003EP-00025638.
PF
XX 08-JUL-1999; 99JP-00194486.
PR 11-JAN-2000; 2000JP-00118774.
PR 02-MAY-2000; 2000JP-00183865.
PR 07-JUL-2000; 2000EP-00114089.
XX
PA (REAS-) RES ASSOC BIOTECHNOLOGY.
XX
PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX
DR WPI; 2004-204755/20.
XX
XX New oligonucleotide primers (830 cDNAs) useful for synthesizing full
PT length human cDNAs.
XX
PS Example 18; SEQ ID NO 2077; 1340pp; English.
XX
CC This invention relates to a novel primers useful for synthesising full
CC length cDNA molecules that encode human proteins. Specifically, it refers
CC to secretory or membrane proteins that are potential therapeutic agents/
CC target molecules in the field of medicine, and in particular genes
CC encoding proteins that are associated with signal transduction,
CC glycoproteins and transcription. The present invention describes a method
CC for efficiently cloning a full length human cDNA from both the 5' and 3'
CC ends using the oligo-capping method. This polynucleotide sequence is the
CC 3' end of a representative human DNA cluster of the invention.
XX
SQ Sequence 825 BP; 160 A; 247 C; 230 G; 183 T; 0 U; 5 Other;

Query Match 99.6%; Score 588.4; DB 12; Length 825;
Best Local Similarity 99.7%; Pred. No. 2e-134;
Matches 589; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGGAGGGAGGCTTGAAGAGGAAACACTCTGATTTGGAAGAGGAGGAGGAGGTGGGAG 60
Db |||||
QY 93 ATGGTGGGAGGCTTGAAGAGGAAACACTCTGATTTGGAAGAGGAGGAGGAGGTGGGAG 152
Db |||||
QY 61 TGGAGTCCAGCAGGCCCTTCAGAGCTACCAAGCAAGCCCTGCTCCGCATCTCCCTAGACAAA 120
Db |||||
QY 153 TGGAGTCCAGCAGGCCCTTCAGAGCTACCAAGCAAGCCCTGCTCCGCATCTCCCTAGACAAA 212
Db |||||
QY 121 GTCCAGCGCAGCCTGGGCCCCCGAGCACCCAGCCTCCGAGGCATGTCCTCATCCATAAC 180
Db |||||
QY 213 GTCCAGCGCAGCCTGGGCCCCCGAGCACCCAGCCTCCGAGGCATGTCCTCATCCATAAC 272
Db |||||
QY 181 ACCCTCCAACAGCTGCAGGTGCACCTTGGCTGGCTCCGCCCTGCCCTGCCCTGCCCTGCCGAG 240
Db |||||

PA (RANK/) RANK D R.
PA (HANK/) HANZEL D K.
XX
PI Penn SG, Rank DR, Hanzel DK;
XX
XX WPI; 2004-119264/12.
DR
XX
PT New human genome-derived single exon nucleic acid probes useful for human
PT gene expression analysis, for identifying or characterizing alternative
PT splicing events, for assessing genomic alterations or as tools for
PT surveying tissues.
XX
PS Claim 1; SEQ ID NO 24898; 80pp; English.
XX
CC The invention relates to a nucleic acid probe for measuring human gene
CC expression, comprising any of the 27,400 fully defined nucleotide
CC sequences in the specification, or their complements or fragments, and
CC encoding at least 8 amino acids of any of the 6888 amino acid sequences
CC fully defined in the specification. The probe is a single exon probe that
CC hybridises under high stringency conditions to a nucleic acid molecule
CC expressed in human cells or tissues. Also included are a spatially-
CC addressable set of single exon nucleic acid probes for measuring human
CC gene expression (comprising a plurality of single exon nucleic acid
CC probes cited above, where each of the plurality of probes is separately
CC and addressably isolatable or amplifiable from the plurality), a single
CC exon microarray for measuring human gene expression, a method of
CC measuring human gene expression, a vector comprising the single exon
CC probe cited above, an ORF-encoded peptide comprising at least 8
CC contiguous amino acids of any of the above-mentioned amino acid
CC sequences (optionally with conservative amino acid substitutions), an
CC isolated antibody that binds specifically to a peptide cited above,
CC methods of selling and/or licensing single exon probes or microarrays to
CC a customer desiring to measure gene expression, a method of providing
CC human gene expression data by subscription, and a computer-readable
CC storage medium which contains a database having a plurality of records
CC (each record including data on the expression of a single exon probe
CC cited above. The probe, methods and apparatus are useful in gene
CC expression analysis. The probes may be used as tools for surveying
CC tissues to detect the presence of expressed messages that contain their
CC specific exon, or in constructing genome-derived single exon microarrays.
CC In addition, the probes are used in identifying and characterising
CC alternative splicing events, in detecting and characterising gross
CC alterations in the genomic locus that includes their exon, in assessing
CC smaller genomic alterations, in priming the synthesis of nucleic acids,
CC or in expressing the ORF-encoded peptide. The present sequence is a human
CC single exon probe of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?DocID=20030194704
XX
SQ Sequence 598 BP; 123 A; 197 C; 158 G; 120 T; 0 U; 0 Other;

Query Match 97.9%; Score 578.4; DB 12; Length 598;
Best Local Similarity 99.7%; Pred. No. 5.3e-132;
Matches 590; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Qy 1 ATGGAGGGAGGCTTGAAGAGGAAACACTCTGATTGTGGAAGAGGAGGAGGAGGTGGGAG 60
Db |||||
7 ATGGTGGGAGGCTTGAAGAGGAAACACTCTGATTGTGGAAGAGGAGGAGGAGGTGGGAG 66

Qy 61 TGGAGTCCAGCAGGCGCTTGCAGAGCTACCAGCAAGCCCTGCTCCGCATCTCCCTAGACAAA 120
Db |||||
67 TGGAGTCCAGCAGGCGCTTGCAGAGCTACCAGCAAGCCCTGCTCCGCATCTCCCTAGACAAA 126

Qy 121 GTCCAGCGCAGCCTGGGCCCCCGAGCACCCAGCCCTCCGAGGCATGTCCTCATCCATAAC 180
Db |||||
127 GTCCAGCGCAGCCTGGGCCCCCGAGCACCCAGCCCTCCGAGGCATGTCCTCATCCATAAC 186

Qy 181 ACCCTCCAACAGCTGCAGCTGCACCTTCGCCTGGCTCCCGCCCCCTGCCCTG-CCCCCGGA 239
Db |||||
187 ACCCTCCAACAGCTGCAGCTGCACCTTCGCCTGGCTTCGCGCCCTGCCCCCCCCCGGA 246

Qy 240 GCCCCCTCTTCTTCTGGGCGAGGAGGATTCTTCCCTGTGAGCCACCATTGGCTCTATCCTCAG 299

Db 247 GCCCCTCTTCTTGGGCGAGGAGGATTCTCCCTGTGAGCCACCATTGGCTCTATCCTCAG 306
Qy 300 GGAGCTGGACACCTCCATGGATGGGACTGAGCCCCCTCAGAATCCAGTACTCCCTTGG 359
Db 307 GGAGCTGGACACCTCCATGGATGGGACTGAGCCCCCTCAGAATCCAGTACTCCCTTGG 366
Qy 360 CCTCCAGAAATGAAGTGCCACCCAGCCTGATCCAGTCTTCTTAGAAGCTCTGAGCTCCCG 419
Db 367 CCTCCAGAAATGAAGTGCCACCCAGCCTGATCCAGTCTTCTTAGAAGCTCTGAGCTCCCG 426
Qy 420 GTACTTGGGGACTCTGGCCTGGATGACTTCTTTCTGGACATTGACACATCTGCGGTAGA 479
Db 427 GTACTTGGGGACTCTGGCCTGGATGACTTCTTTCTGGACATTGACACATCTGCGGTAGA 486
Qy 480 AAAGGAGCCTGCACGGGCCCCCACCAGAGCCTCCTCACAACCTCTTCTGTGCCCCAGGTTT 539
Db 487 AAAGGAGCCTGCACGGGCCCCCACCAGAGCCTCCTCACAACCTCTTCTGTGCCCCAGGTTT 546
Qy 540 TTGGGAGTGGAAATGAACCTGGATCACATCATCATGGAATCATTTCTGGGTCCTAA 591
Db 547 TTGGGAGTGGAAATGAACCTGGATCACATCATCATGGAATCATTTCTGGGTCCTAA 598

RESULT 11
AAH34147
ID AAH34147 standard; cDNA; 1596 BP.
XX
AC AAH34147;
XX
DT 03-SEP-2001 (first entry)
XX
DE Human colon cancer antigen encoding cDNA SEQ ID NO:1229.
XX
KW Human; colon cancer; colon cancer antigen; diagnosis; detection;
XX colorectal carcinoma; ss.
OS Homo sapiens.
PN WO200122920-A2.
XX
PD 05-APR-2001.
XX
PF 28-SEP-2000; 2000WO-US026524.
XX
PR 29-SEP-1999; 99US-0157137P.
PR 03-NOV-1999; 99US-0163280P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Ruben SM, Barash SC, Birse CE, Rosen CA;
XX
DR WPI; 2001-235357/24.
DR P-PSDB; AAG74742.
XX
PT Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
XX useful for preventing, diagnosing and/or treating colorectal cancers.
PS Claim 1; Page 3003-3004; 9803pp; English.
XX
CC AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
CC cancer-associated nucleic acid molecules (N) and proteins (P), where the
CC proteins are collectively known as colon cancer antigens. The colon
CC cancer antigens have cytostatic activity and can be used in gene therapy
CC and vaccine production. N and P may be used in the prevention, diagnosis
CC and treatment of diseases associated with inappropriate P expression. For
CC example, N and P may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of P by expressing inactive proteins or to
CC supplement the patients own production of P. Additionally, N may be used
CC to produce the colon cancer-associated Ps, by inserting the nucleic acids
CC into a host cell and culturing the cell to express the proteins. N and P
CC can be used in the prevention, diagnosis and treatment of colorectal

QY 540 TTGGGAGTGAATGA 554
Db 492 TTGGGAGTGAATGA 506

RESULT 13
ABL83261
ID ABL83261 standard; cDNA; 205 BP.
XX
AC ABL83261;
XX
DT 17-MAY-2002 (first entry)
XX
DE Human ovarian cancer related cDNA clone SEQ ID NO:6239.
XX
KW Human; ovarian cancer; ovarian tumour; cytostatic; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200192581-A2.
XX
PD 06-DEC-2001.
XX
PF 29-MAY-2001; 2001WO-US017756.
XX
PR 26-MAY-2000; 2000US-0207484P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Algate PA, Harlocker SL, Jones R;
XX
DR WPI; 2002-122075/16.
XX
PT Composition for therapy and diagnosis of ovarian cancer comprising
PT polypeptide of a ovarian tumor polypeptide, polynucleotide encoding
PT polypeptide, antibody specific to polypeptide or T cell expressing
PT polypeptide.
XX
PS Claim 1; SEQ ID NO 6239; 489pp; English.
XX

The present invention describes a composition (I) comprising: carriers
and immunostimulants; and a polypeptide (III) of a ovarian tumour
polypeptide encoded by a polynucleotide (III) having a cDNA sequence (S1)
from the 10912 nucleotide sequences as given in ABL77023 to ABL87934,
(III) encoding (II) having a sequence (S2), a T cell population of (II),
or antigen presenting cells that express (II). (I) has cytostatic
activity. An oligonucleotide (IV) that hybridises to (S1) can be used for
detecting ovarian cancer in a patient's biological sample preferably
serum or ovarian tissue. The method comprises contacting a biological
sample from a patient with (IV), detecting the amount of polynucleotide
hybridising to (IV) and comparing the amount to a predetermined cutoff
value and thereby detecting ovarian cancer in the patient, where the
amount of polynucleotide hybridising to (IV) is detected preferably by
polymerase chain reaction (PCR). (I) comprising (III) and/or (II) is
useful for stimulating and/or expanding T cells specific for an ovarian
tumour protein comprising contacting T cells with (III) or (II). (III) is
useful in design and preparation of ribozyme molecules for inhibiting
expression of the tumour polypeptides and proteins in tumour cells; and
to isolate a full length gene from a suitable library e.g., a tumour cDNA
library using well known techniques

Sequence 205 BP; 36 A; 89 C; 45 G; 35 T; 0 U; 0 Other;

Query Match 31.6%; Score 186.6; DB 6; Length 205;
Best Local Similarity 97.1%; Pred. No. 5.2e-36;
Matches 201; Conservative 0; Mismatches 4; Indels 2; Gaps 1;

Db 61 CCAGCGC--CCTGGGCCCCGAGCACCCAGCCTCCGCAGCATTTGCTCATCATAACAC 118
QY 183 CCTCCAACAGCTGCAGGCTGCACCTTCGCCTGGCTCCGCCCTGCCCCCGAGCC 242
Db 119 CCTCCAACAGCTGCAGGCTGCACCTTCGCCTGGCTCCGCCCTGCCCCGAGCC 178
QY 243 CCTCTTCTCTGGCGCAGGAGGATTTCTC 269
Db 179 CCTCTTCTCTGGCGCAGGAGGATTTCTC 205

RESULT 14
ACH77983
ID ACH77983 standard; DNA; 526 BP.
XX
AC ACH77983;
XX
DT 29-JUL-2004 (first entry)
XX
DE Human genome derived single exon probe #11178.
XX
KW Human; probe; ss; gene expression; single exon probe; microarray;
KW alternative splicing event; genomic alteration.
XX
OS Homo sapiens.
XX
PN US2003194704-A1.
XX
PD 16-OCT-2003.
XX
PF 03-APR-2002; 2002US-00029386.
XX
PR 03-APR-2002; 2002US-00029386.
XX
PA (PENN/) PENN S G.
PA (RANK/) RANK D R.
PA (HANZ/) HANZEL D K.
XX
PI Penn SG, Rank DR, Hanzel DK;
XX
DR WPI; 2004-119264/12.
XX
PT New human genome-derived single exon nucleic acid probes useful for human
PT gene expression analysis, for identifying or characterizing alternative
PT splicing events, for assessing genomic alterations or as tools for
PT surveying tissues.
XX
PS Claim 15; SEQ ID NO 11178; 80pp; English.
XX
CC The invention relates to a nucleic acid probe for measuring human gene
CC expression, comprising any of the 27,400 fully defined nucleotide
CC sequences in the specification, or their complements or fragments, and
CC encoding at least 8 amino acids of any of the 6888 amino acid sequences
CC fully defined in the specification. The probe is a single exon probe that
CC hybridises under high stringency conditions to a nucleic acid molecule
CC expressed in human cells or tissues. Also included are a spatially-
CC addressable set of single exon nucleic acid probes for measuring human
CC gene expression (comprising a plurality of single exon nucleic acid
CC probes cited above, where each of the plurality of probes is separately
CC and addressably isolatable or amplifiable from the plurality), a single
CC exon microarray for measuring human gene expression, a method of
CC measuring human gene expression, a vector comprising the single exon
CC probe cited above, an ORF-encoded peptide comprising at least 8
CC contiguous amino acids of any of the above- mentioned amino acid
CC sequences (optionally with conservative amino acid substitutions), an
CC isolated antibody that binds specifically to a peptide cited above,
CC methods of selling and/or licensing single exon probes or microarrays to
CC a customer desiring to measure gene expression, a method of providing
CC human gene expression data by subscription, and a computer-readable
CC storage medium which contains a database having a plurality of records
CC (each record including data on the expression of a single exon probe
CC cited above. The probe, methods and apparatus are useful in gene

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model
Run on: April 26, 2005, 06:54:54 ; Search time 155 Seconds
(without alignments)
6238.965 Million cell updates/sec

Title: US-10-069-386A-1
Perfect score: 591
Sequence: 1 atggaggagggttggaag.....aaatcattctggggctctaa 591

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCITUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	49.2	8.3	7218	1	US-08-232-463-14
2	46.4	7.9	44952	4	US-09-949-016-12197
3	46.4	7.9	44960	4	US-09-949-016-17583
4	46.4	7.9	51552	4	US-09-733-294A-30
5	42.6	7.2	1603	4	US-10-000-489-71
6	41	6.9	1926	3	US-09-249-585A-2
7	41	6.9	1926	4	US-09-410-399-3
8	41	6.9	2580	3	US-09-050-863-2
9	41	6.9	2580	3	US-09-359-081-2
10	41	6.9	5452	2	US-09-130-114-1
11	41	6.9	8705	4	US-09-647-344A-14
12	41	6.9	9600	3	US-08-910-647-1
13	41	6.9	9600	3	US-09-620-925-1
14	41	6.9	10596	1	US-07-884-811-15
15	41	6.9	10596	1	US-07-885-971-15
16	41	6.9	10596	1	US-08-087-783A-15
17	41	6.9	10596	1	US-08-194-088B-15
18	41	6.9	10596	2	US-08-194-087-15
19	41	6.9	10596	5	PCT-US93-04648-15
20	41	6.9	16080	4	US-09-724-566A-48
21	41	6.9	16080	4	US-09-471-669A-48
22	40.8	6.9	50937	3	US-09-428-517-1
23	40.6	6.9	396	1	US-07-872-678A-9
24	40.4	6.8	2097	3	US-08-941-445A-10
25	40.2	6.8	2254	4	US-09-016-434-1126
26	40.2	6.8	2573	2	US-08-884-681-2
27	40.2	6.8	2573	3	US-09-258-643-2

28	40.2	6.8	4403765	3	US-09-103-840A-2	Sequence 2, Appli
29	40.2	6.8	4411529	3	US-09-103-840A-1	Sequence 1, Appli
c 30	40	6.8	2127	4	US-09-902-540-8358	Sequence 8358, Ap
c 31	40	6.8	7628	4	US-09-902-540-875	Sequence 875, App
c 32	39.6	6.7	1432	4	US-09-620-312D-420	Sequence 420, App
c 33	39.6	6.7	2825	3	US-09-196-390-5	Sequence 5, Appli
c 34	39.6	6.7	2825	4	US-09-952-677-5	Sequence 5, Appli
35	39.6	6.7	4394	1	US-08-095-734-1	Sequence 1, Appli
36	39.6	6.7	4394	2	US-08-444-623-1	Sequence 1, Appli
37	39.6	6.7	4394	3	US-08-471-869-1	Sequence 1, Appli
38	39.6	6.7	4394	3	US-09-342-563-1	Sequence 1, Appli
39	39.6	6.7	4394	5	PCT-US94-08267-1	Sequence 1, Appli
40	39.4	6.7	780	4	US-09-902-540-4017	Sequence 4017, Ap
41	39.4	6.7	2481	4	US-09-894-988A-35	Sequence 35, Appl
42	39.4	6.7	2481	4	US-10-237-551-35	Sequence 35, Appl
43	39.4	6.7	3066	4	US-10-237-551-152	Sequence 152, App
c 44	39.4	6.7	26533	4	US-09-902-540-1199	Sequence 1199, Ap
45	39.4	6.7	154746	4	US-09-827-688-8	Sequence 8, Appli

ALIGNMENTS

RESULT 1
US-08-232-463-14
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZgpt-Fls
US-08-232-463-14

Query Match

8.3%; Score 49.2; DB 1; Length 7218;

APPLICANT: William Gaarde
APPLICANT: Susan M. Freier
APPLICANT: Edward V. Wanciewicz
TITLE OF INVENTION: ANTISENSE MODULATION OF TERT EXPRESSION
FILE REFERENCE: ISPH-0527
CURRENT APPLICATION NUMBER: US/09/733,294A
PRIOR FILING DATE: 2000-12-07
PRIOR APPLICATION NUMBER: 09/572,423
PRIOR FILING DATE: 2000-05-16
NUMBER OF SEQ ID NOS: 108
SEQ ID NO 30
LENGTH: 51552
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: exon
LOCATION: (1)...(11492)
OTHER INFORMATION: exon 1
NAME/KEY: intron
LOCATION: (11493)...(11596)
OTHER INFORMATION: intron 1
NAME/KEY: exon
LOCATION: (11597)...(12950)
OTHER INFORMATION: exon 2
NAME/KEY: intron
LOCATION: (12951)...(21566)
OTHER INFORMATION: intron 2
NAME/KEY: exon
LOCATION: (21567)...(21762)
OTHER INFORMATION: exon 3
NAME/KEY: intron
LOCATION: (21763)...(23851)
OTHER INFORMATION: intron 3
NAME/KEY: exon
LOCATION: (23852)...(24032)
OTHER INFORMATION: exon 4
NAME/KEY: intron
LOCATION: (24033)...(24719)
OTHER INFORMATION: intron 4
NAME/KEY: exon
LOCATION: (24720)...(24899)
OTHER INFORMATION: exon 5
NAME/KEY: intron
LOCATION: (24900)...(25393)
OTHER INFORMATION: intron 5
NAME/KEY: exon
LOCATION: (25394)...(25549)
OTHER INFORMATION: exon 6
NAME/KEY: intron
LOCATION: (25550)...(30196)
OTHER INFORMATION: intron 6
NAME/KEY: exon
LOCATION: (30195)...(30292)
OTHER INFORMATION: exon 7
NAME/KEY: intron
LOCATION: (30293)...(31272)
OTHER INFORMATION: intron 7
NAME/KEY: exon
LOCATION: (31273)...(31358)
OTHER INFORMATION: exon 8
NAME/KEY: intron
LOCATION: (31359)...(33843)
OTHER INFORMATION: intron 8
NAME/KEY: unsure
LOCATION: 31450
OTHER INFORMATION: unknown
NAME/KEY: exon
LOCATION: (33844)...(33957)
OTHER INFORMATION: exon 9
NAME/KEY: intron
LOCATION: (33958)...(35941)
OTHER INFORMATION: intron 9
NAME/KEY: exon

LOCATION: (35942)...(36013)
OTHER INFORMATION: exon 10
NAME/KEY: intron
LOCATION: (36014)...(37884)
OTHER INFORMATION: intron 10
NAME/KEY: exon
LOCATION: (37885)...(38073)
OTHER INFORMATION: exon 11
NAME/KEY: intron
LOCATION: (38074)...(41874)
OTHER INFORMATION: intron 11
NAME/KEY: exon
LOCATION: (41875)...(42001)
OTHER INFORMATION: exon 12
NAME/KEY: intron
LOCATION: (42002)...(42881)
OTHER INFORMATION: intron 12
NAME/KEY: exon
LOCATION: (42882)...(42943)
OTHER INFORMATION: exon 13
NAME/KEY: intron
LOCATION: (42944)...(46129)
OTHER INFORMATION: intron 13
NAME/KEY: exon
LOCATION: (46130)...(46254)
OTHER INFORMATION: exon 14
NAME/KEY: intron
LOCATION: (46255)...(47035)
OTHER INFORMATION: intron 14
NAME/KEY: exon
LOCATION: (47036)...(47173)
OTHER INFORMATION: exon 15
NAME/KEY: intron
LOCATION: (47174)...(47709)
OTHER INFORMATION: intron 15
NAME/KEY: exon
LOCATION: (47710)...(50544)
OTHER INFORMATION: exon 16
US-09-733-294A-30

Query Match 7.9%; Score 46.4; DB 4; Length 51552;
Best Local Similarity 50.4%; Pred. No. 0.024;

Matches 139; Conservative 0; Mismatches 136; Indels 1; Gaps 1;

QY 45 GGAGGAGAGTGGAGTGGAGTCCAGCAGGCGCTTACAGAGTACCAGCAAGCCCTGCTCCG 104
Db 48820 GGAGGCAGGGAGGAGGGGCCAGCCAGGCTTCCCATCTTCCCGCCACCCAGACCAG 48761
QY 105 CATCTCCCTAGACAAAGTCCAGCGCAGCCTGGGGCCCCCGAGCACCAGCCTCCGCAGGCA 164
Db 48760 CCCCCCATCACCCCTGTCACCCAGCTGGGGCCCCCATCATCCCTGCCACCTGGCCA 48701
QY 165 TGTCTCATCCATAACACCT-CCAAACAGCTGCAGGCTGCAGTTCGCTGGCTCCCGCC 223
Db 48700 GGCCCTCCATCATCCCGCGCCCAAGCTGGGGCCCCCAGCATCCCTGCTGCCCTTCGGSC 48641
QY 224 CTGCCCTGCCCCCGAGCCCTCTTCTCTGGGGGAGGAGGATTTCTCCCTGTCAGCCACCA 283
Db 48640 CTGGACTTACTGTATTCTTCCAGGGTGGGGGCTCCCACTGTCTATCCCTACCTCCT 48581
QY 284 TTGGCTCTATCCTCAGGAGCTGGACACCTCCATGG 319
Db 48580 TCCCTCCTGCCTCACAGCATCAGAAACCTCCCAGG 48545

RESULT 5
US-10-000-489-71
; Sequence 71, Application US/10000489
; Patent No. 6794363
; GENERAL INFORMATION:
; APPLICANT: Benjanin, Stephane
; APPLICANT: Tanaka, Hiroaki
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF

FILE REFERENCE: 91.US6.DIV
CURRENT APPLICATION NUMBER: US/10/000,489
CURRENT FILING DATE: 2001-11-14
PRIOR APPLICATION NUMBER: US 09/924,340
PRIOR FILING DATE: 2001-08-06
PRIOR APPLICATION NUMBER: PCT/IB01/01715
PRIOR FILING DATE: 2001-08-06
PRIOR APPLICATION NUMBER: US 60/305,456
PRIOR FILING DATE: 2001-07-13
PRIOR APPLICATION NUMBER: US 60/302,277
PRIOR FILING DATE: 2001-06-29
PRIOR APPLICATION NUMBER: US 60/298,698
PRIOR FILING DATE: 2001-06-15
PRIOR APPLICATION NUMBER: US 60/293,574
PRIOR FILING DATE: 2001-05-25
NUMBER OF SEQ ID NOS: 112
SOFTWARE: JPatent
SEQ ID NO 71
LENGTH: 1603
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: 5'UTR
LOCATION: 1..7
NAME/KEY: CDS
LOCATION: 8..763
NAME/KEY: 3'UTR
LOCATION: 764..1603
NAME/KEY: polyA_signal
LOCATION: 1562..1567
NAME/KEY: polyA_site
LOCATION: 1588..1603
US-10-000-489-71

Query Match 7.2%; Score 42.6; DB 4; Length 1603;
Best Local Similarity 54.0%; Pred. No. 0.066;
Matches 87; Conservative 0; Mismatches 74; Indels 0; Gaps 0;
QY 131 GCCTGGGCCCCGAGCACCCAGCCTCCGCGAGGATGTCCTCATCCATAACACCCCTCCAAC 190
Db 870 GCACGGAACCCCGCCCACTCCCAATCCCCGCGCCCTCTCCACACCCGTGCTCCCC 929
QY 191 AGCTGCAGGCTGCATTGCGCTGGCTCCCGCCCTGCGCTGCCCTGCCCCCGAGCCCTCTTCC 250
Db 930 CGCTCCACCCCTCACCTCACCTCGCCCCCGCCCCACCCATCGCGCCCCCGGCGTGTAT 989
QY 251 TGGCGGAGGAGGATTTCTCCTGTGTCAGCCACCATTTGGCTCT 291
Db 990 TGTTCGGTGGGCTCGGTGCGGCGCTGTCTCCCTCGGCTCT 1030

RESULT 6
US-09-249-585A-2/c
Sequence 2, Application US/09249585A
Patent No. 6417002
GENERAL INFORMATION:
APPLICANT: Horlick, Robert
TITLE OF INVENTION: METHOD FOR MAINTENANCE AND SELECTION OF EPISOMES
FILE REFERENCE: 0867/0D905
CURRENT APPLICATION NUMBER: US/09/249,585A
CURRENT FILING DATE: 1999-02-11
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PatentIn version 3.0
SEQ ID NO 2
LENGTH: 1926
TYPE: DNA
ORGANISM: Epstein Barr Virus
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(1926)
OTHER INFORMATION: coding strand of EBNA-1 DNA
US-09-249-585A-2

Query Match 6.9%; Score 41; DB 3; Length 1926;
Best Local Similarity 51.4%; Pred. No. 0.2;
Matches 95; Conservative 0; Mismatches 90; Indels 0; Gaps 0;
QY 67 CCAGCAGGCCTTCAGAGCTACCAGCAAGCCTGCTCCGCACTCTCCCTAGACAAAGTCCAG 126
Db 1063 CCCGGCGGCTCCACTACCTCCTCGACCCCGGCTCCACTACCTCCTCGACCCCGGCTC 1004
QY 127 CGCAGCCTGGGCCCCGAGCACCCAGCCTCCGCGAGGCATGTCTCATCCATAAACCCCTC 186
Db 1003 CACTGCCTCCTCGACCCCGGCTCCACCTCCTGCTCCTGCCCCCTCCTGCTCCTGCTCCTC 944
QY 187 CAACAGCTGCAGGCTGCACCTTGGCCTGGCTCCGCGCCCTGCCCCCTGCCCCGAGCCCTC 246
Db 943 CTCCTGCTCCTGCCCCCTCCTGCCCCCTCCTGCTCCTGCCCCCTCCTGCTCCTGCTCCTG 884
QY 247 TTCCT 251
Db 883 CCCCT 879

RESULT 7
US-09-410-399-3/c
Sequence 3, Application US/09410399
Patent No. 6482587
GENERAL INFORMATION:
APPLICANT: Robertson, Erle S.
APPLICANT: Cotter, Murray A.
TITLE OF INVENTION: Methods to Inhibit or Enhance the Binding of Viral DNA
TITLE OF INVENTION: to Genomic Host DNA
FILE REFERENCE: UM-03778
CURRENT APPLICATION NUMBER: US/09/410,399
CURRENT FILING DATE: 1999-10-01
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 3
LENGTH: 1926
TYPE: DNA
ORGANISM: Epstein-Barr virus
US-09-410-399-3

Query Match 6.9%; Score 41; DB 4; Length 1926;
Best Local Similarity 51.4%; Pred. No. 0.2;
Matches 95; Conservative 0; Mismatches 90; Indels 0; Gaps 0;
QY 67 CCAGCAGGCCTTCAGAGCTACCAGCAAGCCTGCTCCGCACTCTCCCTAGACAAAGTCCAG 126
Db 1063 CCCGGCGGCTCCACTACCTCCTCGACCCCGGCTCCACTACCTCCTCGACCCCGGCTC 1004
QY 127 CGCAGCCTGGGCCCCGAGCACCCAGCCTCCGCGAGGCATGTCTCATCCATAAACCCCTC 186
Db 1003 CACTGCCTCCTCGACCCCGGCTCCACCTCCTGCTCCTGCCCCCTCCTGCTCCTGCTCCTC 944
QY 187 CAACAGCTGCAGGCTGCACCTTGGCCTGGCTCCGCGCCCTGCCCCCTGCCCCGAGCCCTC 246
Db 943 CTCCTGCTCCTGCCCCCTCCTGCCCCCTCCTGCTCCTGCCCCCTCCTGCTCCTGCTCCTG 884
QY 247 TTCCT 251
Db 883 CCCCT 879

RESULT 8
US-09-050-863-2/c
Sequence 2, Application US/09050863
Patent No. 6114111
GENERAL INFORMATION:
APPLICANT: Lao, Ying
APPLICANT: Hiang, Betty
APPLICANT: Payan, Don
TITLE OF INVENTION: Mammalian Protein Interaction Cloning
SYSTEM
NUMBER OF SEQUENCES: 5

;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert
;; STREET: 4 Embarcadero Center, Suite 3400
;; CITY: San Francisco
;; STATE: CA
;; COUNTRY: USA
;; ZIP: 94111-4187
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/050,863
;; FILING DATE: 30-MAR-1998
;; CLASSIFICATION:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Silva, Robin M.
;; REGISTRATION NUMBER: 38,304
;; REFERENCE/DOCKET NUMBER: A-65638/DJB/RMS
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (415) 781-1989
;; TELEFAX: (415) 949-8711
;; INFORMATION FOR SEQ ID NO: 2:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 2580 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: unknown
;; TOPOLOGY: DNA
;; MOLECULE TYPE: DNA
US-09-050-863-2

Query Match 6.9%; Score 41; DB 3; Length 2580;
Best Local Similarity 51.4%; Pred. No. 0.22;
Matches 95; Conservative 0; Mismatches 90; Indels 0; Gaps 0;

QY 67 CCAGCAGGCCTTCAGAGCTACCAAGCCCTGCTCCGCATCTCCCTAGACAAAGTCCAG 126
Db 1446 CCGGGCGCCTCCACTACCTCCTCGACCCCGGCTCCACTACCTCCTCGACCCCGGCTC 1387

QY 127 GCGAGCCTGGGCCCCGAGCACCAGCCCTCGCAGGATGCTCTCATCCATAACACCCTC 186
Db 1386 CACTGCCTCCTCGACCCCGGCTCCACCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1327

QY 187 CAACAGCTGCAGGCTGCACTTGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTC 246
Db 1326 CTCCTGCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1267

QY 247 TTCCT 251
Db 1266 CCCCT 1262

RESULT 9
US-09-359-081-2/c
; Sequence 2, Application US/09359081
; Patent No. 6316223
; GENERAL INFORMATION:
; APPLICANT: Lao, Ying
; Hiang, Betty
; Payan, Don
; TITLE OF INVENTION: Mammalian Protein Interaction Cloning
; SYSTEM
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/359,081
;; FILING DATE: 22-Jul-1999
;; CLASSIFICATION: <Unknown>
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 09/050,863
;; FILING DATE: <Unknown>
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Silva, Robin M.
;; REGISTRATION NUMBER: 38,304
;; REFERENCE/DOCKET NUMBER: A-65638/DJB/RMS
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (415) 781-1989
;; TELEFAX: (415) 949-8711
;; INFORMATION FOR SEQ ID NO: 2:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 2580 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: unknown
;; TOPOLOGY: unknown
;; MOLECULE TYPE: DNA
;; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-359-081-2

Query Match 6.9%; Score 41; DB 3; Length 2580;
Best Local Similarity 51.4%; Pred. No. 0.22;
Matches 95; Conservative 0; Mismatches 90; Indels 0; Gaps 0;

QY 67 CCAGCAGGCCTTCAGAGCTACCAAGCCCTGCTCCGCATCTCCCTAGACAAAGTCCAG 126
Db 1446 CCGGGCGCCTCCACTACCTCCTCGACCCCGGCTCCACTACCTCCTCGACCCCGGCTC 1387

QY 127 GCGAGCCTGGGCCCCGAGCACCAGCCCTCGCAGGATGCTCTCATCCATAACACCCTC 186
Db 1386 CACTGCCTCCTCGACCCCGGCTCCACCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1327

QY 187 CAACAGCTGCAGGCTGCACTTGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTC 246
Db 1326 CTCCTGCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1267

QY 247 TTCCT 251
Db 1266 CCCCT 1262

RESULT 10
US-09-130-114-1
; Sequence 1, Application US/09130114
; Patent No. 5976807
; GENERAL INFORMATION:
; APPLICANT: Horlick, Robert A.
; APPLICANT: Damaj, Bassam B.
; APPLICANT: Robbins, Alan K.
; TITLE OF INVENTION: Eukaryotic Cells Stably Expressing Genes
; TITLE OF INVENTION: From Multiple Transfected Episomes
; FILE REFERENCE: 0867/ID903US1
; CURRENT APPLICATION NUMBER: US/09/130,114
; CURRENT FILING DATE: 1998-08-06
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 5452
; TYPE: DNA
; ORGANISM: VEBNA
US-09-130-114-1

Query Match 6.9%; Score 41; DB 2; Length 5452;
Best Local Similarity 51.4%; Pred. No. 0.31;
Matches 95; Conservative 0; Mismatches 90; Indels 0; Gaps 0;

;
;
; FILING DATE: 21-Jul-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/910,647
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Fujita, Sharon M.
; REGISTRATION NUMBER: 38,459
; REFERENCE/DOCKET NUMBER: 1218.002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 923-2706
; TELEFAX: (510) 655-3542
;
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9600 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-620-925-1

Query Match 6.9%; Score 41; DB 3; Length 9600;
Best Local Similarity 51.4%; Pred. No. 0.39;
Matches 95; Conservative 0; Mismatches 90; Indels 0; Gaps 0;

QY 67 CCAGCAGGCGCTTCAGAGCTACCAAGCCCTGCTCCGCGCATCTCCCTAGACAAAGTCCAG 126
Db 1492 CCCGGCGGCTCCACTACCTCCTCGACCCCGGCTCCACTACCTCCTCGACCCGGCCTC 1433

QY 127 CGCAGCCTGGGCCCCGAGCAGCAGCCCTCCGCGCGCATGTCTCATCCATAACACCCCTC 186
Db 1432 CACTGCCTCCTCGACCCGGCCTCCACCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1373

QY 187 CAACAGCTGCAGGTGCACCTGCGCTGGCTCCGCGCGCTGCTGCTGCTGCTGCTGCTGCTG 246
Db 1372 CTCCTGCTCCTGCCCCCTCCTGCCCCCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1313

QY 247 TTCCT 251
Db 1312 CCCCT 1308

RESULT 14
US-07-884-811-15/c
; Sequence 15, Application US/07884811
; Patent No. 5316921
; GENERAL INFORMATION:
; APPLICANT: Godowski, Paul J. Lokker, Nathalie A. Mark, Melanie R.
; TITLE OF INVENTION: SINGLE-CHAIN HEPATOCYTE GROWTH FACTOR VARIANTS
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/884,811
; FILING DATE: 19920518
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Dreger, Ginger R.
; REGISTRATION NUMBER: 33,055

;
;
; REFERENCE/DOCKET NUMBER: 755.1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-3216
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
;
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10596 bases
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-07-884-811-15

Query Match 6.9%; Score 41; DB 1; Length 10596;
Best Local Similarity 51.4%; Pred. No. 0.4;
Matches 95; Conservative 0; Mismatches 90; Indels 0; Gaps 0;

QY 67 CCAGCAGGCGCTTCAGAGCTACCAAGCCCTGCTCCGCGCATCTCCCTAGACAAAGTCCAG 126
Db 3026 CCCGGCGGCTCCACTACCTCCTCGACCCCGGCTCCACTACCTCCTCGACCCCGGCTC 2967

QY 127 CGCAGCCTGGGCCCCGAGCAGCAGCCCTCCGCGCGCATGTCTCATCCATAACACCCCTC 186
Db 2966 CACTGCCTCCTCGACCCCGGCTCCACCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2907

QY 187 CAACAGCTGCAGGTGCACCTGCGCTGGCTCCGCGCGCTGCTGCTGCTGCTGCTGCTGCTG 246
Db 2906 CTCCTGCTCCTGCCCCCTCCTGCCCCCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2847

QY 247 TTCCT 251
Db 2846 CCCCT 2842

RESULT 15
US-07-885-971-15/c
; Sequence 15, Application US/07885971
; Patent No. 5328837
; GENERAL INFORMATION:
; APPLICANT: Godowski, Paul J. Lokker, Nathalie A. Mark, Melanie R.
; TITLE OF INVENTION: HEPATOCYTE GROWTH FACTOR PROTEASE DOMAIN VARIANTS
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/885,971
; FILING DATE: 19920518
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Dreger, Ginger R.
; REGISTRATION NUMBER: 33,055
; REFERENCE/DOCKET NUMBER: 779
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-3216
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
;
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10596 bases
; TYPE: NUCLEIC ACID

```

; STRANDEDNESS: single
; TOPOLOGY: linear
US-07-885-971-15

Query Match      6.9%; Score 41; DB 1; Length 10596;
Best Local Similarity 51.4%; Pred. No. 0.4;
Matches 95; Conservative 0; Mismatches 90; Indels 0; Gaps 0;

QY      67 CCAGCAGGCCTTCAGAGCTACCAGCAAGCCCTGCTCCGCATCTCCCTAGACAAAAGTCCAG 126
Db      67 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY      127 CGCAGCCTGGGCCCCCGAGACACCCAGCCCTCCGCGAGGCATGTCTCATCCATAACACCCCTC 186
Db      127 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY      2966 CACTGCCTCCTCGACCCCGGCCTCCACCTCCTGTCTCCTGCCCCCTCCTGTCTCCTGCCCCCTC 2907
Db      2966 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY      187 CAACAGCTGCAGGCTGCACCTTCGCCCTGGCTCCCGCCCCCTGCCCTGCCCCCCCGAGCCCCCTC 246
Db      187 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY      2906 CTCCTGTCTCCTGCCCCCTCCTGTGCCCCCTCCTGTGCCCCCTCCTGTGCCCCCTCCTGTGCTCCTG 2847
Db      2906 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY      247 TTCCT 251
Db      247 |||
QY      2846 CCCCT 2842
Db      2846 CCCCT 2842
```

Search completed: April 26, 2005, 10:14:01
Job time : 161 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 26, 2005, 09:13:38 ; Search time 500 Seconds
(without alignments)
7175.983 Million cell updates/sec

Title: US-10-069-386A-1
Perfect score: 591
Sequence: 1 atggaggaggcttgaagag.....aaatcattctggggtcctaa 591

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 5633728 seqs, 303525691 residues

Total number of hits satisfying chosen parameters: 11267456

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*
19: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
20: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
21: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
22: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	578.4	97.9	598	16	US-10-029-386-24898 Sequence 24898, A
2	576.2	97.5	1596	15	US-10-106-698-1239 Sequence 1239, Ap
3	473	80.0	506	9	US-09-867-701-6064 Sequence 6064, Ap
4	186.6	31.6	205	9	US-09-867-701-6239 Sequence 6239, Ap
5	166	28.1	526	16	US-10-029-386-11178 Sequence 11178, A
6	47.4	8.0	2082	13	US-10-076-069-1 Sequence 1, Appli
7	46.4	7.9	3179	19	US-10-840-455-20 Sequence 20, Appl
8	46.4	7.9	25138	19	US-10-840-455-44 Sequence 44, Appl
9	46.4	7.9	51552	9	US-09-733-294A-30 Sequence 30, Appl
10	44.8	7.6	440	14	US-10-184-644-202 Sequence 202, App
11	44.8	7.6	440	14	US-10-184-634-202 Sequence 202, App

C 12	44.8	7.6	440	16	US-10-063-685-52	Sequence 52, Appl
C 13	44.8	7.6	594	14	US-10-123-155-10	Sequence 10, Appl
C 14	44.8	7.6	594	15	US-10-146-731-10	Sequence 10, Appl
C 15	44.8	7.6	594	15	US-10-140-472-10	Sequence 10, Appl
C 16	44.8	7.6	594	15	US-10-141-761-10	Sequence 10, Appl
C 17	44.8	7.6	594	16	US-10-142-885-10	Sequence 10, Appl
C 18	44.8	7.6	594	16	US-10-158-790-10	Sequence 10, Appl
C 19	44.8	7.6	594	17	US-10-137-871-10	Sequence 10, Appl
C 20	44.8	7.6	594	17	US-10-140-923-10	Sequence 10, Appl
C 21	44.8	7.6	594	17	US-10-141-756-10	Sequence 10, Appl
C 22	44.8	7.6	594	17	US-10-141-759-10	Sequence 10, Appl
C 23	44.8	7.6	594	17	US-10-140-805-10	Sequence 10, Appl
C 24	44.8	7.6	594	17	US-10-140-864-10	Sequence 10, Appl
C 25	44.8	7.6	594	17	US-10-142-426-10	Sequence 10, Appl
26	44.8	7.6	2748	18	US-10-437-963-53188	Sequence 53188, A
27	42.6	7.2	1603	10	US-09-992-600A-71	Sequence 71, Appl
28	42.6	7.2	1603	10	US-09-924-340-71	Sequence 71, Appl
29	42.6	7.2	1603	10	US-09-992-095B-71	Sequence 71, Appl
30	42.6	7.2	1603	10	US-09-999-570-71	Sequence 71, Appl
31	42.6	7.2	1603	14	US-10-000-489-71	Sequence 71, Appl
32	42.6	7.2	1603	14	US-10-000-986-71	Sequence 71, Appl
33	42.6	7.2	1603	16	US-10-154-678-71	Sequence 71, Appl
34	42.6	7.2	1603	16	US-10-001-142-71	Sequence 71, Appl
35	42.6	7.2	1603	19	US-10-838-854-71	Sequence 71, Appl
36	42.4	7.2	511	16	US-10-029-386-24981	Sequence 24981, A
C 37	42.4	7.2	600	18	US-10-363-345A-39087	Sequence 39087, A
38	42.4	7.2	600	18	US-10-363-345A-39088	Sequence 39088, A
C 39	42.4	7.2	600	19	US-10-363-483A-39087	Sequence 39087, A
40	42.4	7.2	600	19	US-10-363-483A-39088	Sequence 39088, A
C 41	42.4	7.2	1650	18	US-10-437-963-27296	Sequence 27296, A
42	42	7.1	672	18	US-10-437-963-69519	Sequence 69519, A
43	42	7.1	2580	15	US-10-156-761-3895	Sequence 3895, Ap
44	42	7.1	9025608	15	US-10-156-761-1	Sequence 1, Appli
45	41.8	7.1	32329	17	US-10-374-903A-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1

US-10-029-386-24898
; Sequence 24898, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 24898
; LENGTH: 598
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC010271.5
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.64
; OTHER INFORMATION: EST HUMAN HIT: BG720189.1, EVALUE 0.00e+00
; OTHER INFORMATION: SWISSPROT HIT: O70622, EVALUE 2.30e-01
; OTHER INFORMATION: NT HIT: gil16178011, EVALUE 0.00e+00
US-10-029-386-24898

Query Match 97.9%; Score 578.4; DB 16; Length 598;
Best Local Similarity 99.7%; Pred. No. 3e-161;
Matches 590; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Qy 1 ATGGAGGGAGGCTTGAAGAGGAACACTCTGATTTGGAAGAGGAGGAGGAGGTGGGAG 60
|||||

Db 7 ATGGTGGGAGGCTTGAAGAGGAACACTCTGATTTGGAAGAGGAGGAGGAGGTGGGAG 66
|||||

QY 61 TGGAGTCCAGCAGGCCTTCAGAGCTACCAGCAAGCCCTGCTCCGCATCTCCCTAGACAAA 120
Db 67 TGGAGTCCAGCAGGCCTTCAGAGCTACCAGCAAGCCCTGCTCCGCATCTCCCTAGACAAA 126
QY 121 GTCCAGCGCAGCCTGGGCCCCCGAGCAGCACCAGCCCTCCGAGGCATGTCTCATCCATAAC 180
Db 127 GTCCAGCGCAGCCTGGGCCCCCGAGCAGCACCAGCCCTCCGAGGCATGTCTCATCCATAAC 186
QY 181 ACCCTCCAACAGCTGCAGGTGCACATTTCGCCCTGGCTCCGCCCCCTGCCCTG-CCCCCGA 239
Db 187 ACCCTCCAACAGCTGCAGGTGCACATTTCGCCCTGGCTCCGCCCCCTGCCCTGCCCCCGCA 246
QY 240 GCGCCTCTTCTGGCGAGGAGGATTTCCTCCCTGTCAAGCCACCATTGGCTCTATCTCAG 299
Db 247 GCGCCTCTTCTGGCGAGGAGGATTTCCTCCCTGTCAAGCCACCATTGGCTCTATCTCAG 306
QY 300 GGAGCTGGACACCTCCATGGATGGGACTGAGCCCCCTCAGAAATCCAGTACTCCCTTGG 359
Db 307 GGAGCTGGACACCTCCATGGATGGGACTGAGCCCCCTCAGAAATCCAGTACTCCCTTGG 366
QY 360 CCTCCAGAATGAAGTGCCACCCAGCCTGATCCAGTCTTCTTAGAAGCTCTGAGTCCCCG 419
Db 367 CCTCCAGAATGAAGTGCCACCCAGCCTGATCCAGTCTTCTTAGAAGCTCTGAGTCCCCG 426
QY 420 GTACTTGGGGGACTCTGGCCTGGATGACTTCTTCTTGACATTGACACATCTGCGGTAGA 479
Db 427 GTACTTGGGGGACTCTGGCCTGGATGACTTCTTCTTGACATTGACACATCTGCGGTAGA 486
QY 480 AAGGAGCCTGCACGGGCCCCACAGAGCCTCTCTACAACCTTCTTGTGCCCCAGGTTT 539
Db 487 AAGGAGCCTGCACGGGCCCCACAGAGCCTCTCTACAACCTTCTTGTGCCCCAGGTTT 546
QY 540 TTGGGAGTGAATGAACCTGGATCACATCATGGAATCATTTCTGGGGTCCCTAA 591
Db 547 TTGGGAGTGAATGAACCTGGATCACATCATGGAATCATTTCTGGGGTCCCTAA 598

RESULT 2
US-10-106-698-1239
; Sequence 1239, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide
; FILE REFERENCE: PA005P1
; CURRENT APPLICATION NUMBER: US/10/106,698
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 1239
; LENGTH: 1596
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (57)..(57)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-106-698-1239

Query Match 97.5%; Score 576.2; DB 15; Length 1596;
Best Local Similarity 99.3%; Pred. No. 1.5e-160;
Matches 587; Conservative 2; Mismatches 0; Indels 2; Gaps 1;
QY 1 ATGGAGGGAGGCTTGAAGAGGAAACACTCTGATTTTGAAGAGGAGGAGGAGGAGGTTGGGAG 60
Db 345 ATGGAGGGAGGCTTGAAGAGGAAACACTCTGATTTTGAAGAGGAGGAGGAGGAGGTTGGGAG 404

QY 61 TGGAGTCCAGCAGGCCTTCAGAGCTACCAGCAAGCCCTGCTCCGCATCTCCCTAGACAAA 120
Db 405 TGGAGTCCAGCAGGCCTTCAGAGCTACCAGCAAGCCCTGCTCCGCATCTCCCTAGACAAA 464
QY 121 GTCCAGCGCAGCCTGGGCCCCCGAGCAGCACCAGCCCTCCGAGGCATGTCTCATCCATAAC 180
Db 465 GTCCAGCGC--CCTGGGCCCCCGAGCAGCACCAGCCCTCCGAGGCATGTCTCATCCATAAC 522
QY 181 ACCCTCCAACAGCTGCAGGTGCACATTTCGCCCTGGCTCCGCCCCCTGCCCTGCCCTCCGAG 240
Db 523 ACCCTCCAACAGCTGCAGGTGCACATTTCGCCCTGGCTCCGCCCCCTGCCCTGCCCTCCGAG 582
QY 241 CCCCTCTTCTGGCGAGGAGGATTTCCTCCCTGTCAAGCCACCATTGGCTCTATCTCAGG 300
Db 583 CCCCTCTTCTGGCGAGGAGGATTTCCTCCCTGTCAAGCCAMCATTTGGCTCTATCTCAGG 642
QY 301 GAGCTGGACACCTCCATGGATGGGACTGAGCCCCCTCAGAAATCCAGTACTCCCTTGGC 360
Db 643 GAGCTGGACACCTCCATGGATGGGACTGAGCCCCCTCAGAAATCCAGTACTCCCTTGGC 702
QY 361 CTCCAGAATGAAGTGCCACCCAGCCTGATCCAGTCTTCTTAGAAGCTCTGAGTCCCCGG 420
Db 703 CTCCAGAATGAAGTGCCACCCAGCCTGATCCAGTCTTCTTAGAAGCTCTGAGTCCCCGG 762
QY 421 TACTTGGGGGACTCTGGCCTGGATGACTTCTTCTTGACATTGACACATCTGCGGTAGAA 480
Db 763 TACTTGGGGGACTCTGGCCTGGATGACTTCTTCTTGACATTGACACATCTGCGGTAGAA 822
QY 481 AAGGAGCCTGCACGGGCCCCACAGAGCCTCTCTACAACCTTCTTGTGCCCCAGGTTTCT 540
Db 823 AAGGAGCCTGCACGGGCCCCACAGAGCCTTCTCTACAACCTTCTTGTGCCCCAGGTTTCT 882
QY 541 TGGGAGTGAATGAACCTGGATCACATCATGGAATCATTTCTGGGGTCCCTAA 591
Db 883 TGGGAGTGAATGAACCTGGATCACATCATGGAATCATTTCTGGGGTCCCTAA 933

RESULT 3
US-09-867-701-6064
; Sequence 6064, Application US/09867701
; Patent No. US20020132237A1
; GENERAL INFORMATION:
; APPLICANT: Aglate, Paul A.
; APPLICANT: Jones, Robert
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.497
; CURRENT APPLICATION NUMBER: US/09/867,701
; CURRENT FILING DATE: 2001-05-29
; NUMBER OF SEQ ID NOS: 10912
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6064
; LENGTH: 506
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-867-701-6064

Query Match 80.0%; Score 473; DB 9; Length 506;
Best Local Similarity 98.6%; Pred. No. 5.1e-130;
Matches 488; Conservative 0; Mismatches 5; Indels 2; Gaps 1;
QY 60 GTGGAGTCCAGCAGGCCTTCAGAGCTACCAGCAAGCCCTGCTCCGCATCTCCCTAGACAA 119
Db 14 GCGGAGTCCAGCAGGCCTTCAGAGCTACCAGCAAGCCCTGCTCCGCATCTCCCTAGACAA 73
QY 120 AGTCCAGCGCAGCCTGGGCCCCCGAGCAGCACCAGCCTCCGAGGCATGTCTCATCCATAA 179
Db 74 AGTCCAGCGC--CCTGGGCCCCCGAGCAGCACCAGCCTCCGAGCAGTGTCTCATCCATAA 131
QY 180 CACCCTCCAACAGCTGCAGGTGCACATTTCGCCCTGGCTCCGCCCCCTGCCCTGCCCTCCCGA 239

Db 132 CACCCTCCAACAGCTGCAGGATGCACTTCGCCTGGCTCCGGCCCCCTGCCCTGCCCCCCCGA 191

QY 240 GCCCTCTTCCTGGCGGAGGAGGATTTCTCCCTGTCTCAGCCACCATTTGGCTCTATCCTCAG 299

Db 192 GCCCTCTTCCTGGCGGAGGAGGATTTCTCCCTGTCTCAGCCACCATTTGGCTCTATCCTCAG 251

QY 300 GGAGCTGGACACCTCCATGGATGGGACTGAGCCCCCTCAGAAATCCAGTGACTCCCTTGG 359

Db 252 GGAGCTGGACACCTCCATGGATGGGACTGAGCCCCCTCAGAAATCCAGTGACTCCCTTGG 311

QY 360 CCTCCAGAATGAAGTGCCACCCAGCCTGATCCAGTCTTCTTAGAAGCTCTGAGTCCCG 419

Db 312 CCTCCAGAATGAAGTGCCACCCAGCCTGATCCAGTCTTCTTAGAAGCTCTGAGTCCCG 371

QY 420 GTACTTGGGGACTCTGGCCTGGATGACTTCTTTCTGGACATTGACACATCTGCGGTAGA 479

Db 372 GTACTTGGGGACTCTGGCCTGGATGACTTCTTTCTGGACATTGACACATCTGCGGTAGA 431

QY 480 AAAGGAGCCTGCACGGGCCCCCACCAGAGCCTCCTCACAACTCTTCTGTGCCCAAGTTC 539

Db 432 AAAGGAGCCTGCACGGGCCCCCACCAGAGCCTCCTCACAACTCTTCTGTGCCCAAGTTC 491

QY 540 TTGGGAGTGAATGA 554

Db 492 TTGGGAGTGAATGA 506

RESULT 4

US-09-867-701-6239

; Sequence 6239, Application US/09867701

; Patent No. US20020132237A1

; GENERAL INFORMATION:

; APPLICANT: Aglate, Paul A.

; APPLICANT: Jones, Robert

; APPLICANT: Harlocker, Susan L.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY

; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER

; FILE REFERENCE: 210121.497

; CURRENT APPLICATION NUMBER: US/09/867,701

; CURRENT FILING DATE: 2001-05-29

; NUMBER OF SEQ ID NOS: 10912

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 6239

; LENGTH: 205

; TYPE: DNA

; ORGANISM: Homo sapien

US-09-867-701-6239

Query Match 31.6%; Score 186.6; DB 9; Length 205;

Best Local Similarity 97.1%; Pred. No. 3.5e-45;

Matches 201; Conservative 0; Mismatches 4; Indels 2; Gaps 1;

QY 63 GAGTCCAGCAGCGCCTTCAGAGCTACCAGCAAGCCCTGCTCCGCATCTCCCTAGACAAAGT 122

Db 1 GAGTCCAGCAGCGCCTTCAGAGCTACCAGCAAGCCCTGCTCCGCATCTCCCTAGACAAAGT 60

QY 123 CCAGCGCAGCCTGGGCCCCCGAGCACCCAGCCTCCGAGGCATGTCCTCATCCATAACAC 182

Db 61 CCAGCGC--CCTGGGCCCCCGAGCACCCAGCCTCCGAGCATTGTCTCATCCATAACAC 118

QY 183 CCTCCAACAGCTGCAGGCTGCACTTGCGCTGGCTCCGCGCCCTGCCCTGCCCCCGAGGCC 242

Db 119 CCTCCAACAGCTGCAGGCTGCACTTGCGCTGGCTCCGCGCCCTGCCCTGCACCCCGAGGCC 178

QY 243 CTTCTTCTCTGGCGAGGAGGATTTCTC 269

Db 179 CTTCTTCTCTGGCGAGGAGGATTTCTC 205

RESULT 5

US-10-029-386-11178

; Sequence 11178, Application US/10029386

; Publication No. US20030194704A1

; GENERAL INFORMATION:

; APPLICANT: Penn, Sharron G.

; APPLICANT: Rank, David R.

; APPLICANT: Hanzel, David K.

; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR GI

; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO

; FILE REFERENCE: AEOMICA-X-2

; CURRENT APPLICATION NUMBER: US/10/029,386

; CURRENT FILING DATE: 2001-12-20

; NUMBER OF SEQ ID NOS: 34288

; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1

; SEQ ID NO 11178

; LENGTH: 526

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; OTHER INFORMATION: MAP TO AC010271.5

; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.64

; OTHER INFORMATION: SWISSPROT HIT: O60963, EVALUE 2.20e+00

; OTHER INFORMATION: NT HIT: g115718683, EVALUE 0.00e+00

; OTHER INFORMATION: EST_HUMAN HIT: AA426355.1, EVALUE 1.00e-114

US-10-029-386-11178

Query Match 28.1%; Score 166; DB 16; Length 526;

Best Local Similarity 100.0%; Pred. No. 5.2e-39;

Matches 166; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 426 GGGGGACTCTGGCCTGGATGACTTCTTTCTGGACATTGACACATCTGCGGTAGAAAAGGA 485

Db 1 GGGGGACTCTGGCCTGGATGACTTCTTTCTGGACATTGACACATCTGCGGTAGAAAAGGA 60

QY 486 GCCTGCACGGGCCCCCACCAGAGCCTCCTCACAACTTCTGTGCCCCAGGTTCTTTGGGA 545

Db 61 GCCTGCACGGGCCCCCACCAGAGCCTCCTCACAACTTCTGTGCCCCAGGTTCTTTGGGA 120

QY 546 GTGGAATGAACCTGGATCACATCATGGAATCATTTGGGGTCCTAA 591

Db 121 GTGGAATGAACCTGGATCACATCATGGAATCATTTGGGGTCCTAA 166

RESULT 6

US-10-076-069-1

; Sequence 1, Application US/10076069

; Publication No. US20020177214A1

; GENERAL INFORMATION:

; APPLICANT: JURECIC, ROLAND

; APPLICANT: NACHTMAN, RONALD

; TITLE OF INVENTION: HEPP, A NOVEL GENE WITH A ROLE IN HEMATOPOIETIC AND NEURAL DEVELOP

; FILE REFERENCE: 39532-176599

; CURRENT APPLICATION NUMBER: US/10/076,069

; CURRENT FILING DATE: 2002-02-15

; PRIOR APPLICATION NUMBER: US 60/268,923

; PRIOR FILING DATE: 2001-02-16

; NUMBER OF SEQ ID NOS: 11

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 1

; LENGTH: 2082

; TYPE: DNA

; ORGANISM: Mus musculus

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (191)..(901)

US-10-076-069-1

Query Match 8.0%; Score 47.4; DB 13; Length 2082;

Best Local Similarity 55.0%; Pred. No. 0.0009;

Matches 115; Conservative 0; Mismatches 91; Indels 3; Gaps 1;

QY 8 GAGGCTTGAAGAGGAAACACTCTGATTTGGAAGAGGAGGAGGAGGTGGAGTGGAGTC 67

Db 201 GAGGCTGAAGAGGAAATATGGTGACCAGGAAGAGGAGTAGAGGTTTGT---GCACGTG 257

QY 68 CAGCAGGCCTTCAGAGCTACCAGCAAGCCCTGCTCCGCATCTCCCTAGACAAAGTCCAGC 127


```
Db      258 TCCCTTCCTATAGCCTGCAGCGACAGTCACTCTCTGGACATGTCCCTTGTCAGCTCCAGC 317
QY      128 GCAGCCTGGGCCCCCGAGCACCAGCCCTCCGCGAGGCATGTCCTCATCCATAACACCCCTCC 187
Db      318 TCTGTACATGCTAGTGGAGCCCAATCTCTGCCGTCGGTCTCTCATCGCCAAACAGTCC 377
QY      188 AACAGCTGCAGGCTGCACCTTCGCCTGGCT 216
Db      378 GGCAGATCCAGGAGGAAATGAGCCAGGAT 406

RESULT 7
US-10-840-455-20/c
; Sequence 20, Application US/10840455
; Publication No. US20050032094A1
; GENERAL INFORMATION:
; APPLICANT: Bayer Aktiengesellschaft
; APPLICANT: Hagen, Gustav
; APPLICANT: Wick, Maresa
; APPLICANT: Zubov, Dmitry
; TITLE OF INVENTION: Regulatory DNA Sequences of the Gene for the Human Catalytic
; TITLE OF INVENTION: Telomerase Subunit, and Their Diagnostic and Therapeutic Use
; FILE REFERENCE: Lea 32 805C1
; CURRENT APPLICATION NUMBER: US/10/840,455
; PRIOR FILING DATE: 2004-05-06
; PRIOR APPLICATION NUMBER: PCT/EP98/08216
; PRIOR FILING DATE: 1998-12-22
; PRIOR APPLICATION NUMBER: US 09/582,246
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: DE19757984.1
; PRIOR FILING DATE: 1997-12-24
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 20
; LENGTH: 3179
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-840-455-20
```

```
Query Match      7.9%; Score 46.4; DB 19; Length 3179;
Best Local Similarity 50.4%; Pred. No. 0.0019;
Matches 139; Conservative 0; Mismatches 136; Indels 1; Gaps 1;

QY      45 GGAGGAGAGGTGGGAGTGGAGTCCAGCAGGCCCTTCAGAGCTACCCAGCAAGCCCTGCTCCG 104
Db      447 GGAGGCAGGGGAGGAGGGGGCCCCAGCCAGGCTTCCCCATCTTCCCCGCCACCCAGACCAG 388
QY      105 CATCTCCCTAGACAAAGTCCAGCGCAGCCTGGGCCCCCGAGCACCCAGCCCTCCGCGAGGCA 164
Db      387 CCCCCCCCATCACCCCTGCCACCCAGCTGGGGCCCCCATCATCCCTGCCACCCCTGGCCA 328
QY      165 TGTCTCTCATCCATAACACCCT-CCAACAGCTGCAGGCTGCACCTTCGCCCTGGCTCCCGCCC 223
Db      327 GGCCCTCCATCATCCCGCGCCGCCAAGCTGGGCCCCCAGCATCCCTGTGCTGCCCTTCGGGC 268
QY      224 CTGCCCTGCCCCCCCGAGCCCCCTCTTCTCTGGGCGAGGAGGATTTCTCCCTGTGACGCCACCA 283
Db      267 CTGGACTTACTGTTATGTTCTTCCAGGGTGGGGCTCCCACTGTCTATCCCCCTACCCCTCCT 208
QY      284 TTGGCTCTATCCTCAGGGAGCTGGACACCTCCATGG 319
Db      207 TCCCTCTCGCCTCACAGCATCAGAAACCTCCCAAG 172
```

```
RESULT 8
US-10-840-455-44/c
; Sequence 44, Application US/10840455
; Publication No. US20050032094A1
; GENERAL INFORMATION:
; APPLICANT: Bayer Aktiengesellschaft
; APPLICANT: Hagen, Gustav
; APPLICANT: Wick, Maresa
```

```
; APPLICANT: Zubov, Dmitry
; TITLE OF INVENTION: Regulatory DNA Sequences of the Gene for the Human Catalytic
; TITLE OF INVENTION: Telomerase Subunit, and Their Diagnostic and Therapeutic Use
; FILE REFERENCE: Lea 32 805C1
; CURRENT APPLICATION NUMBER: US/10/840,455
; CURRENT FILING DATE: 2004-05-06
; PRIOR APPLICATION NUMBER: PCT/EP98/08216
; PRIOR FILING DATE: 1998-12-22
; PRIOR APPLICATION NUMBER: US 09/582,246
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: DE19757984.1
; PRIOR FILING DATE: 1997-12-24
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 44
; LENGTH: 25138
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(25138)
; OTHER INFORMATION: n is a, c, g, or t.
US-10-840-455-44

Query Match      7.9%; Score 46.4; DB 19; Length 25138;
Best Local Similarity 50.4%; Pred. No. 0.0025;
Matches 139; Conservative 0; Mismatches 136; Indels 1; Gaps 1;

QY      45 GGAGGAGAGGTGGGAGTGGAGTCCAGCAGGCCCTTCAGAGCTACCCAGCAAGCCCTGCTCCG 104
Db      22406 GGAGGCAGGGGAGGAGGGGGCCCCAGCCAGGCTTCCCCATCTTCCCCGCCACCCAGACCAG 22347
QY      105 CATCTCCCTAGACAAAGTCCAGCGCAGCCTGGGCCCCCGAGCACCCAGCCCTCCGCGAGGCA 164
Db      22346 CCCCCCCCATCACCCCTGCCACCCAGCTGGGGCCCCCATCATCCCTGCCACCCCTGGCCA 22287
QY      165 TGTCTCTCATCCATAACACCCT-CCAACAGCTGCAGGCTGCACCTTCGCCCTGGCTCCCGCCC 223
Db      22286 GGCCCTCCATCATCCCGCGCCGCCAAGCTGGGGCCCCCAGCATCCCTGTGCTGCCCTTCGGGC 22227
QY      224 CTGCCCTGCCCCCCCGAGCCCCCTCTTCTCTGGGCGAGGAGGATTTCTCCCTGTGACGCCACCA 283
Db      22226 CTGGACTTACTGTTATGTTCTTCCAGGGTGGGGCTCCCACTGTCTATCCCCCTACCCCTCCT 22167
QY      284 TTGGCTCTATCCTCAGGGAGCTGGACACCTCCATGG 319
Db      22166 TCCCTCTCGCCTCACAGCATCAGAAACCTCCCAAG 22131
```

```
RESULT 9
US-09-733-294A-30/c
; Sequence 30, Application US/09733294A
; Patent No. US20020045588A1
; GENERAL INFORMATION:
; APPLICANT: Brett P. Monia
; APPLICANT: William Gaarde
; APPLICANT: Susan M. Freier
; APPLICANT: Edward V. Wanciewicz
; TITLE OF INVENTION: ANTISENSE MODULATION OF TERT EXPRESSION
; FILE REFERENCE: ISPH-0527
; CURRENT APPLICATION NUMBER: US/09/733,294A
; CURRENT FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: 09/572,423
; PRIOR FILING DATE: 2000-05-16
; NUMBER OF SEQ ID NOS: 108
; SEQ ID NO 30
; LENGTH: 51552
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: exon
; LOCATION: (1)...(11492)
; OTHER INFORMATION: exon 1
```


NAME/KEY: intron
LOCATION: (11493)...(11596)
OTHER INFORMATION: intron 1
NAME/KEY: exon
LOCATION: (11597)...(12950)
OTHER INFORMATION: exon 2
NAME/KEY: intron
LOCATION: (12951)...(21566)
OTHER INFORMATION: intron 2
NAME/KEY: exon
LOCATION: (21567)...(21762)
OTHER INFORMATION: exon 3
NAME/KEY: intron
LOCATION: (21763)...(23851)
OTHER INFORMATION: intron 3
NAME/KEY: exon
LOCATION: (23852)...(24032)
OTHER INFORMATION: exon 4
NAME/KEY: intron
LOCATION: (24033)...(24719)
OTHER INFORMATION: intron 4
NAME/KEY: exon
LOCATION: (24720)...(24899)
OTHER INFORMATION: exon 5
NAME/KEY: intron
LOCATION: (24900)...(25393)
OTHER INFORMATION: intron 5
NAME/KEY: exon
LOCATION: (25394)...(25549)
OTHER INFORMATION: exon 6
NAME/KEY: intron
LOCATION: (25550)...(30196)
OTHER INFORMATION: intron 6
NAME/KEY: exon
LOCATION: (30195)...(30292)
OTHER INFORMATION: exon 7
NAME/KEY: intron
LOCATION: (30293)...(31272)
OTHER INFORMATION: intron 7
NAME/KEY: exon
LOCATION: (31273)...(31358)
OTHER INFORMATION: exon 8
NAME/KEY: intron
LOCATION: (31359)...(33843)
OTHER INFORMATION: intron 8
NAME/KEY: unsure
LOCATION: 31450
OTHER INFORMATION: unknown
NAME/KEY: exon
LOCATION: (33844)...(33957)
OTHER INFORMATION: exon 9
NAME/KEY: intron
LOCATION: (33958)...(35941)
OTHER INFORMATION: intron 9
NAME/KEY: exon
LOCATION: (35942)...(36013)
OTHER INFORMATION: exon 10
NAME/KEY: intron
LOCATION: (36014)...(37884)
OTHER INFORMATION: intron 10
NAME/KEY: exon
LOCATION: (37885)...(38073)
OTHER INFORMATION: exon 11
NAME/KEY: intron
LOCATION: (38074)...(41874)
OTHER INFORMATION: intron 11
NAME/KEY: exon
LOCATION: (41875)...(42001)
OTHER INFORMATION: exon 12
NAME/KEY: intron
LOCATION: (42002)...(42881)
OTHER INFORMATION: intron 12
NAME/KEY: exon

LOCATION: (42882)...(42943)
OTHER INFORMATION: exon 13
NAME/KEY: intron
LOCATION: (42944)...(46129)
OTHER INFORMATION: intron 13
NAME/KEY: exon
LOCATION: (46130)...(46254)
OTHER INFORMATION: exon 14
NAME/KEY: intron
LOCATION: (46255)...(47035)
OTHER INFORMATION: intron 14
NAME/KEY: exon
LOCATION: (47036)...(47173)
OTHER INFORMATION: exon 15
NAME/KEY: intron
LOCATION: (47174)...(47709)
OTHER INFORMATION: intron 15
NAME/KEY: exon
LOCATION: (47710)...(50544)
OTHER INFORMATION: exon 16
US-09-733-294A-30
Query Match 7.9%; Score 46.4; DB 9; Length 51552;
Best Local Similarity 50.4%; Pred. No. 0.0028;
Matches 139; Conservative 0; Mismatches 136; Indels 1; Gaps 1;
QY 45 GGAGGAGAGTGGGAGTCCAGCAGGCGCTTCCAGAGCTACAGCAAGCCCTGCTCCG 104
Db 48820 GGAGGAGGAGGAGGAGGCGCCAGCCAGGCTTCCCATCTTCCCGCCACCCAGACCAG 48761
QY 105 CATCTCCCTAGACAAAGTCCAGCGCAGCTGGGCGGCGCCAGCAGCAGCCAGCTCCGAGGCA 164
Db 48760 CCCCCCATCATCCCTGCCACCCAGCTGGGGCGCCCATCATCCCTGCCACCTGGGCA 48701
QY 165 TGTCTCATCCATAACACCCCT-CCAAACAGCTGCAGGCTGCACCTTCGCCTGGCTCCGCCC 223
Db 48700 GGCCCTCCATCATCCCGCGCCCAAGCTGGGCGCCAGCATCCCTGCTGCCCTTCGGGC 48641
QY 224 CTGCCCTGCCCCCGAGCCCTCTTCTCTGGGCGAGGAGGATTTCTCCCTGTTCAGCCACCA 283
Db 48640 CTGGACTTACTGTATGTCTTCCAGGTGGGGTCCCACTGTCTATCCCTACCTCTCT 48581
QY 284 TTGGCTCTATCTCAGGAGCTGGACACCTCCATGG 319
Db 48580 TCCCCTCTCTCAGCATCATCAGCAACCTCCAGG 48545

RESULT 10
US-10-184-644-202/c
Sequence 202, Application US/10184644
Publication No. US20030044930A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
TITLE OF INVENTION: ACIDS ENCODING THE SAME
FILE REFERENCE: P3430RIC227
CURRENT APPLICATION NUMBER: US/10/184,644
CURRENT FILING DATE: 2002-06-28
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 612
SEQ ID NO 202
LENGTH: 440
TYPE: PRT


```
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Sherwood, Steven  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Colin K  
; APPLICANT: Wood, William  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P3330R1C323  
; CURRENT APPLICATION NUMBER: US/10/146,731  
; PRIOR FILING DATE: 2002-05-15  
; Prior Application removed - See File Wrapper or Palm  
; NUMBER OF SEQ ID NOS: 550  
; SEQ ID NO 10  
; LENGTH: 594  
; TYPE: PRP  
; ORGANISM: Homo Sapiens  
  
US-10-146-731-10
```

Query Match 7.6% ; Score 44.8 ; DB 15 ; Length 594;
Best Local Similarity 8.7%; Pred. No. 0.0044;
Matches 28; Conservative 135; Mismatches 160; Indels 0; Gaps 0

```
QY   79 CAGAGCTACCGCAAGCCCTGCTCCGCATCTCCCCTAGACAAGAAGTCCAGCGCAGCCTGGGC 138  
      :: : |:: : ||:: : |:|: : : : : : : : : : : : : : : : : :  
Db   556 MM.M.CBT.STT.YA.M.YT.S.S.S.SYSYSYS.S.S.SDSYSYA.SSYS.S.SWS 497  
  
QY   139 CCCGAGCACGCCGTCCGAGGCATGTTCCTCATCCATAAACACCCTCCAACAGCTGCAG 198  
      :::: : : : |:: : |:: : |:: : |:: : |:: : |:: : |:: : |:  
Db   496 SYSYSSDDY.CYCYYRYHCHSDSYSYYY.CRCCYT.TSRYDCHYSCCSDDYYCY 437  
  
QY   199 GCTGCATTTCCTTGCTCCGCTCCGCCCCCTGCCCTGCCCGAGCCCTTTCTGGGCGAG 258  
      :::: : : : : : : : |:: : |:: : |:: : |:: : |:: : |:: : |:  
Db   436 YSYRRYYYSYSYSWISYSYTDVCSYRRCYYSYSSSYSSAYSTSSSSSSSS 377  
  
QY   259 GAGGATTTCCTCCCTGTGAGCCACCATTTGGCTCTATCCTCAGGGAGCTGGACACTCCATG 318  
      |:: : |:: : |:: : |:: : |:: : |:: : |:: : |:: : |:: : |:  
Db   376 YYYSTNYC.T.CC....T..MCAAABCSTTTTTTTTTT.HSCC.SA...A.M..YC.A.S 317  
  
QY   319 GATGGGACTGAGCCCCCTCAGAATCCAGTAGCTCCCTTGGCCTCCAGAATGAAGTGCCA 378  
      :: : : : : : : : : : : : : : : : : : : : : : : : : : : :  
Db   316 YYSYS.SSS.S.SYM.RRA.SHYYTRS..S.MYCY.YM.Y..YY.YYSYYCSRKTM.. 257  
  
QY   379 CCCGAGCCTGATCCAGTCTTCT 401  
      : : : | : : : : : : : : : : : : : : : : : : : : : : :  
Db   256 ..TWDM.T.T..MHMY.KVB.H 234
```

RESULT 15
US-10-140-472-10/c
; Sequence 10, Application US/10140472
; Publication No. US2003013888A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.

```
; APPLICANT: Tumas,Daniel
; APPLICANT: Watanabe,Colin K
; APPLICANT: Wood,William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C168
; CURRENT APPLICATION NUMBER: US/10/140,472
; CURRENT FILING DATE: 2002-05-06
; Prior Apploication removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 10
; LENGTH: 594
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-140-472-10
```

```
Query Match          7.6%; Score 44.8; DB 15; Length 594;
Best Local Similarity 8.7%; Pred. No. 0.0044;
Matches 28; Conservative 135; Mismatches 160; Indels 0; Gaps 0;

QY      79 CAGAGCTACGAGCCCTGCTCCGCATCTCCCTAGACAAAGTCCAGCGCAGCCTGGGC 138
      : : | : : | : : : : : : : : : : : : : : : : : : : : : : :
Db      556 MM.M.CBT.STT.VA.M.YT.S.S.S.SYSYSYS.S.S.SDSYSYA.SYSYS.S.SWS 497

QY      139 CCCCAGACACCCAGCCTCCGAGGCATGTCTCTATCCATAACACCCCTCCAACAGCTGCAG 198
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      496 SYSYSSDDY.CYCCYRYHCDSDSYSYY.CRCCYTT.SYSRYDCHYSCCCSDYYCYS 437

QY      199 GCTGCACCTTGGCTCCGCCCTGCCCTGCCCCCGAGCCCCCTCTTCCTGGGCGAG 258
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      436 YSYSRYYSYSYSWSYSYTDYCSYRCCCYYSYSSSYSSSYSSAYSTSSSSSSSS 377

QY      259 GAGGATTCTCCCTGTCAGCCACCATTGGCTCTATCCTCAGGGAGCTGGACACCTCCATG 318
      | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db      376 YYTSTNYC.T.CC....T..MCAABCSTTTTTTTT..HSCC.SA..A.M..YC.A.S 317

QY      319 GATGGGACTGAGCCCCCCTCAGAAATCCAGTGACTCCCTTGGCCTCCAGAAATGAAGTGCCA 378
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      316 YSYSYS.SSS.S.SYMR.HRA.SHYYTRS..S.MYCY.YM.Y..YY.YYSYCSRKTM... 257

QY      379 CCCCAGCCTGATCCAGTCTTCTT 401
      : : : | : : : : : :
Db      256 ..TMTDM.T.T..MHMY.KYB.H 234
```

Search completed: April 26, 2005, 11:38:08
Job time : 505 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 25, 2005, 12:22:40 ; Search time 175 Seconds
(without alignments)
573.529 Million cell updates/sec

Title: US-10-069-386A-2
Perfect score: 1033
Sequence: 1 MEGGLKRKHSDLSEEEERWE.....APGSWEWNELDHIMEILGS 196

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_03:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1033	100.0	196	2	Q9UJW9	Q9ujw9 homo sapien
2	1026	99.3	196	2	Q96CQ2	Q96cq2 homo sapien
3	871	84.3	197	2	Q9ERC3	Q9erc3 mus musculu
4	217	21.0	236	1	STD1_MOUSE	Q9jl10 mus musculu
5	213	20.6	236	1	STD1_HUMAN	Q9uhv2 homo sapien
6	208	20.1	236	2	Q6P771	Q6p771 rattus norv
7	166	16.1	244	2	Q9DC22	Q9dcz2 mus musculu
8	140.5	13.6	309	1	STD2_MOUSE	Q9jjg5 mus musculu
9	136.5	13.2	248	2	Q6GM81	Q6gm81 xenopus lae
10	136	13.2	237	1	CCA4_MOUSE	Q9cwm2 mus musculu
11	133	12.9	361	2	Q6NXD9	Q6nxd9 brachydanio
12	133	12.9	383	2	Q7ZZ27	Q7zz27 brachydanio
13	125	12.1	314	1	STD2_HUMAN	Q14140 homo sapien
14	115	11.1	1194	2	Q76M68	Q76m68 rattus norv
15	113.5	11.0	911	2	Q80TJ8	Q80tj8 mus musculu
16	112.5	10.9	246	2	Q6DFF1	Q6dff1 xenopus lae
17	105.5	10.2	620	2	Q9BQ18	Q9bq18 homo sapien
18	105.5	10.2	760	2	Q76N32	Q76n32 homo sapien
19	105	10.2	241	1	CCA4_HUMAN	Q9bxi8 homo sapien
20	104.5	10.1	740	2	Q9UPP2	Q9upp2 homo sapien
21	101.5	9.8	728	1	P85B_HUMAN	Q00459 homo sapien
22	100	9.7	846	2	Q6NVČ9	Q6nvc9 brachydanio
23	98.5	9.5	314	2	Q6NXL2	Q6nxl2 mus musculu
24	98	9.5	517	2	Q9EEF8	Q9eef8 anticarsia
25	98	9.5	562	2	Q8NAF0	Q8naf0 homo sapien
26	97.5	9.4	309	2	Q8UZB4	Q8uzb4 grapevine f
27	97	9.4	803	2	O74402	O74402 schizosacch
28	97	9.4	1201	2	Q9C6S1	Q9c6s1 arabidopsis
29	96	9.3	817	1	VRP1_YEAST	P37370 saccharomyc
30	96	9.3	817	2	Q07229	Q07229 saccharomyc
31	95.5	9.2	432	2	Q8BN29	Q8bn29 mus musculu

32	95.5	9.2	446	1	TFE3_MOUSE	Q64092 mus musculu
33	95.5	9.2	572	2	Q7TNC1	Q7tncl mus musculu
34	95	9.2	911	2	Q8CGD1	Q8cgd1 mus musculu
35	95	9.2	917	2	Q8CFQ8	Q8cfq8 mus musculu
36	95	9.2	1092	2	Q8CFT2	Q8cft2 mus musculu
37	94.5	9.1	339	2	Q8PIE4	Q8pie4 xanthomonas
38	94.5	9.1	689	2	Q871U4	Q871u4 neurospora
39	94.5	9.1	1468	2	O70641	O70641 simian t-ly
40	94	9.1	433	2	Q08833	Q08833 human t-cel
41	94	9.1	433	2	Q76QA3	Q76qa3 human t-cel
42	94	9.1	433	2	Q82343	Q82343 human t-cel
43	94	9.1	433	2	Q82440	Q82440 human t-cel
44	94	9.1	433	2	Q9DKS2	Q9dks2 human t-cel
45	94	9.1	433	2	Q9WI14	Q9wi14 human t-cel

ALIGNMENTS

RESULT 1
Q9UJW9 ID Q9UJW9 PRELIMINARY; PRT; 196 AA.
AC Q9UJW9;
DT 01-MAY-2000 (TremBLrel. 13, Created)
DT 01-MAY-2000 (TremBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TremBLrel. 26, Last annotation update)
DE RPA-binding trans-activator.
GN Name=RBT1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20440390; PubMed=10982866; DOI=10.1093/nar/28.18.3478;
RA Cho J.M., Song D.J., Alaoui-Jamali M.A.;
RT "RBT1, a novel transcriptional co-activator, binds the second subunit of Replication Protein A";
RL Nucleic Acids Res. 28:3478-3485 (2000).
DR EMBL; AF192529; AAF05761.1; --
DR InterPro; IPR009263; SERTA.
DR Pfam; PF06031; SERTA; 1.
SQ SEQUENCE 196 AA; 21798 MW; 334333EF7F8A9EBA7 CRC64;

Query Match 100.0%; Score 1033; DB 2; Length 196;
Best Local Similarity 100.0%; Pred. No. 1.7e-71;
Matches 196; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MEGGLKRKHSDLSEEEERWEWS	PAGLQSYQQALLRISLDK	VQSRSLGPRAPSLRRHVL	IHN 60
Db	1	MEGGLKRKHSDLSEEEERWEWS	PAGLQSYQQALLRISLDK	VQSRSLGPRAPSLRRHVL	IHN 60
QY	61	TLQQLQAALRLAPALPPEPL	FLGEEDFSLSATIGSILREL	DTSMGTEPPQNPVTPLG	120
Db	61	TLQQLQAALRLAPALPPEPL	FLGEEDFSLSATIGSILREL	DTSMGTEPPQNPVTPLG	120
QY	121	LQNEVPPQDPVFLEALSSRYL	GDSGLDDFFLDIDTSAVE	KEPARAPPEPPHNLFCAPGS	180
Db	121	LQNEVPPQDPVFLEALSSRYL	GDSGLDDFFLDIDTSAVE	KEPARAPPEPPHNLFCAPGS	180
QY	181	WEWNELDHIMEILGS	196		
Db	181	WEWNELDHIMEILGS	196		

RESULT 2
Q96CQ2 ID Q96CQ2 PRELIMINARY; PRT; 196 AA.
AC Q96CQ2;
DT 01-DEC-2001 (TremBLrel. 19, Created)
DT 01-DEC-2001 (TremBLrel. 19, Last sequence update)
DT 25-OCT-2004 (TremBLrel. 28, Last annotation update)
DE RPA-binding trans-activator.

DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)

DE MGC82098 protein.

GN Name=MGC82098;

OS Xenopus laevis (African clawed frog).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;

OC Xenopodinae; Xenopus.

OX NCBI_TaxID=8355;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Kidney;

RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,

RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,

RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,

RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,

RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,

RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,

RA Whiting M., Touchman J.W., Green E.D., Dickinson M.C.,

RA Blakesley R.W., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,

RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,

RA Jones S.J., Marra M.A.;

RT "Generation and initial analysis of more than 15,000 full-length human

RT and mouse cDNA sequences.";

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).

RN [2]

RP SEQUENCE FROM N.A.

RC TISSUE=Kidney;

RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;

RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,

RA Richardson P.;

RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus

RT initiative.";

RL Dev. Dym. 225:384-391 (2002).

RN [3]

RP SEQUENCE FROM N.A.

RC TISSUE=Kidney;

RA Klein S., Gerhard D.S.;

RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.

DR EMBL; BC074195; AAH74195.1; -

DR InterPro; IPR009263; SERTA.

DR Pfam; PF06031; SERTA; 1.

SQ SEQUENCE 248 AA; 27319 MW; E703D7E9105DF08C CRC64;

Query Match 13.2%; Score 136.5; DB 2; Length 248;

Best Local Similarity 30.9%; Pred. No. 0.012;

Matches 58; Conservative 25; Mismatches 58; Indels 47; Gaps 8;

QY 17 ERWESPAGLQSYQQALLRISLDKQVRSGLGPRAPSLRRHVLIHNTLQQLQAALRL---AP 73

Db ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||

Db 10 ERECCSIPAQSH--CLMNISLVKLRSLRHRVPEPLRHFVIVANTLRRLOGNLQVEQCAP 67

QY 74 -----APALPPEPLFLGEEPDFLSATIGSILRELDTSMGTEPPQ 113

Db ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||

Db 68 DMWKTSECTRSLVVPESKPALENTEDPSSMDASLYSSISTILEDLN-NFEGLS--S 124

QY 114 NPVTPLGLQNEVPPQDPVFLEA-----LSSRY-----LGDSGLDDFFLDID 155

Db :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|

Db 125 SPLPQIEDDQLCAPKANPVGSAEDMVKLASSSSLLSSSPYLLGENLGDN-LIEDIFEDID 183

QY 156 TSAVEKEP 163

Db || :| :|

Db 184 TSMYSDP 191

RESULT 10

CCA4_MOUSE

ID CCA4_MOUSE STANDARD; PRT; 237 AA.

AC Q9CWM2; Q921E8; Q99MP6;

DT 10-OCT-2003 (Rel. 42, Created)

DT 10-OCT-2003 (Rel. 42, Last sequence update)

DT 05-JUL-2004 (Rel. 44, Last annotation update)

DE Cell division cycle associated protein 4 (Hematopoietic progenitor protein).

DE Name=Cdca4; Synonyms=HEPP;

GN Mus musculus (Mouse).

OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Bone marrow;

RX MEDLINE=21375891; PubMed=11482882; DOI=10.1006/bcmd.2001.0434;

RA Abdullah J.M., Jing X., Spassov D.S., Nachtman R.G., Jurecic R.;

RT "Cloning and characterization of Hepp, a novel gene expressed preferentially in hematopoietic progenitors and mature blood cells.";

RL Blood Cells Mol. Dis. 27:667-676 (2001).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Mesonephros;

RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;

RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,

RA Nikaido I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,

RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,

RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,

RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,

RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,

RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,

RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,

RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,

RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,

RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,

RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,

RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,

RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,

RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,

RA Sandelin A., Schneider C., Sempole C.A., Setou M., Shimada K.,

RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,

RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,

RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,

RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,

RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,

RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,

RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,

RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,

RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,

RA Birney E., Hayashizaki Y.;

RT "Analysis of the mouse transcriptome based on functional annotation of

RT 60,770 full-length cDNAs.";

RL Nature 420:563-573 (2002).

RN [3]

RP SEQUENCE OF 13-237 FROM N.A.

RC TISSUE=Breast tumor;

RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,

RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,

RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,

RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,

RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,

RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,

RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,

RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Bouffard G.G.,

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Bouffard G.G.,

